

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES
PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum
Internationales Büro



(43) Internationales Veröffentlichungsdatum
25. Juli 2002 (25.07.2002)

PCT

(10) Internationale Veröffentlichungsnummer
WO 02/057464 A2

(51) Internationale Patentklassifikation⁷: C12N 15/82 (74) Anwalt: **GOLDSCHEID**, Bettina; BASF Aktiengesellschaft, 67056 Ludwigshafen (DE).

(21) Internationales Aktenzeichen: PCT/EP02/00461 (81) Bestimmungsstaaten (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, IIR, IIU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(22) Internationales Anmeldedatum: 18. Januar 2002 (18.01.2002) (84) Bestimmungsstaaten (*regional*): ARIPO-Patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI-Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

(25) Einreichungssprache: Deutsch (72) Erfinder; und

(26) Veröffentlichungssprache: Deutsch (75) Erfinder/Anmelder (*nur für US*): **LERCHL**, Jens [DE/SE]; Onsjövägen 17, S-26831 Svalöv (SE). **DUWENIG**, Elke [DE/DE]; Schanzstraße 46, 67063 Ludwigshafen (DE). **BISCHOFF**, Friedrich [DE/DE]; Albinistrasse 11, 55116 Mainz (DE). **HEINZ**, Ernst [DE/DE]; Püttkampsweg 13, 22609 Hamburg (DE). **DREXLER**, Hjördis [DE/DE]; Mendelsohnstr. 33, 22761 Hamburg (DE). **SCHEFFLER**, Jodi [US/US]; 51 County Road 228, Oxford, MS 38655 (US).

(72) Veröffentlicht:
— ohne internationales Recherchenbericht und erneut zu veröffentlichen nach Erhalt des Berichts

Zur Erklärung der Zweibuchstaben-Codes und der anderen Abkürzungen wird auf die Erklärungen ("Guidance Notes on Codes and Abbreviations") am Anfang jeder regulären Ausgabe der PCT-Gazette verwiesen.

(54) Title: METHOD FOR THE EXPRESSION OF BIOSYNTHETIC GENES IN PLANT SEEDS USING NOVEL MULTIPLE EXPRESSION CONSTRUCTS

(54) Bezeichnung: VERFAHREN ZUR EXPRESSION VON BIOSYNTHESEGENEN IN PFLANZLICHEN SAMEN UNTER VERWENDUNG VON NEUEN MULTIPLEN EXPRESSIONSKONSTRUKTEN

(57) Abstract: The invention relates to expression cassettes, combinations thereof and vectors containing said expression cassettes, containing plant promoters with an expression specificity for plant seeds, in particular linseed and the use of said expression cassettes or vectors for the recombinant expression of heterologous genes in plants. The invention further relates to transgenic plants, transformed by means of said expression cassettes, or vectors, cultures, parts or transgenic propagations derived therefrom and the use of the above as foodstuff, animal feedstuff, seedstuff, pharmaceuticals, fine chemicals or industrial raw material.

(57) Zusammenfassung: Die Erfindung betrifft Expressionskassetten, deren Kombination und Vektoren enthaltend die Expressionskassetten, die pflanzliche Promotoren mit einer Expressionsspezifität für pflanzliche Samen insbesondere Leinsamen enthalten, sowie die Verwendung dieser Expressionskassetten oder Vektoren zur rekombinanten Expression heterologer Gene in Pflanzen. Die Erfindung betrifft ferner mit diesen Expressionskassetten oder Vektoren transformierte transgene Pflanzen, davon abgeleitete Kulturen, Teile oder transgenes Vermehrungsgut, sowie die Verwendung derselben als Nahrungs-, Futtermittel oder Saatgut, Pharmazeutika, Feinchemikalien oder als industrieller Grundstoff.

WO 02/057464 A2

Verfahren zur Expression von Biosynthesegenen in pflanzlichen Samen unter Verwendung von neuen multiplen Expressionskonstrukten

5 Beschreibung

Die Erfindung betrifft Expressionskassetten, deren Kombination und Vektoren enthaltend die Expressionskassetten, die pflanzliche Promotoren mit einer Expressionsspezifität für pflanzliche Samen insbesondere Leinsamen enthalten, sowie die Verwendung dieser Expressionskassetten oder Vektoren zur rekombinanten Expression heterologer Gene in Pflanzen. Die Erfindung betrifft ferner mit diesen Expressionskassetten oder Vektoren transformierte transgene Pflanzen, davon abgeleitete Kulturen, Teile oder transgenes Vermehrungsgut, sowie die Verwendung derselben als Nahrungs-, Futtermittel oder Saatgut, Pharmazeutika, Feinchemikalien oder als industrieller Grundstoff.

Die erfindungsgemäßen Expressionskassetten werden vorteilhaft in einem Verfahren zur Herstellung von ungesättigten Fettsäuren mit mindestens zwei Doppelbindungen und/oder ein Verfahren zur Herstellung von Triglyceriden mit erhöhtem Gehalt an mehrfach ungesättigten Fettsäuren mit mindestens zwei Doppelbindungen verwendet. Im Verfahren finden vorteilhaft die Nukleinsäuresequenzen SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder 11 Anwendung, die unter Verwendung der Expressionskassetten exprimiert werden. Diese vorgenannten Nukleinsäuren sind im Verfahren sowie zur Herstellung eines transgenen Organismus bevorzugt einer transgenen Pflanze oder eines transgenen Mikroorganismus mit erhöhtem Gehalt an Fettsäuren, Ölen oder Lipiden mit ungesättigten C₁₈-, C₂₀- oder C₂₂-Fettsäuren geeignet. Außerdem können mit Hilfe der erfindungsgemäßen Expressionskassetten weitere Gene neben den Nukleinäuresequenzen SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 und SEQ ID NO: 11 oder seine Homologen, Derivate oder Analoga sowie Genkonstrukte, die diese Gene oder ihre Homologen, Derivate oder Analoga umfasst, sowie Ihre Verwendung allein oder in Kombination mit weiteren Biosynthesegenen bevorzugt Biosynthesegene für polyungesättigter Fettsäuren, wie vorteilhaft in SEQ ID NO: 7 und SEQ ID NO: 9 dargestellt, in Organismen bevorzugt Pflanzen exprimiert werden.

Eine Reihe von Produkten und Nebenprodukten natürlich vor kommender Stoffwechselprozesse in Mikroorganismen, tierischen und pflanzlichen Zellen sind für viele Industriezweige, einschließlich der Futtermittel-, Nahrungsmittel-, Kosmetik- und pharmazeutischen Industrie, nützlich. Zu diesen gemeinsam als "Feinchemikalien" bezeichneten Molekülen gehören beispielsweise

Lipide und Fettsäuren, unter denen eine beispielhafte Klasse die mehrfach ungesättigten Fettsäuren sind. Fettsäuren und Triglyceride haben eine Vielzahl von Anwendungen in der Lebensmittelindustrie, der Tierernährung, der Kosmetik und im Pharama-
5 bereich. Je nachdem ob es sich um freie gesättigte oder ungesättigte Fettsäuren oder um Triglyceride mit einem erhöhten Gehalt an gesättigten oder ungesättigten Fettsäuren handelt, sind sie für die unterschiedlichsten Anwendungen geeignet, so werden beispielsweise mehrfach ungesättigte Fettsäuren (polyunsaturated fatty acids, PUFAs) Babynahrung zur Erhöhung des Nährwertes zugesetzt. PUFAs haben weiterhin einen positiven Einfluss auf den Cholesterinspiegel im Blut von Menschen und eignen sich daher zum Schutz gegen Herzkrankheiten. So finden sie in verschiedenen diätischen Lebensmitteln oder Medikamenten Anwendung.

15 Besonders geeignete Mikroorganismen zur Herstellung von PUFAs sind Mikroorganismen wie Thraustochytrien oder Schizochytrien-Stämme, Algen wie Phaeodactylum tricornutum oder Cryptocodonium-Arten, Ciliaten, wie Stylonychia oder Colpidium, Pilze, wie Mortierella, Entomophthora oder Mucor. Durch Stammselektion ist eine Anzahl von Mutantenstämmen der entsprechenden Mikro-organismen entwickelt worden, die eine Reihe wünschenswerter Verbindungen, einschließlich PUFAs, produzieren. Die Selektion von Stämmen mit verbesserter Produktion eines bestimmten Moleküls
20 ist jedoch ein zeitraubendes und schwieriges Verfahren.
25

Alternativ kann die Produktion von Feinchemikalien geeigneterweise über die Produktion von Pflanzen, die so entwickelt sind, dass sie die vorstehend genannten PUFAs herstellen, im großen Maßstab durchgeführt werden. Besonders gut für diesen Zweck geeignete Pflanzen sind Ölfruchtpflanzen, die große Mengen an Lipidverbindungen enthalten wie Raps, Canola, Lein, Soja, Sonnenblumen, Borretsch und Nachtkerze. Aber auch andere Nutzpflanzen, die Öle oder Lipide und Fettsäuren enthalten, sind gut geeignet,
30 wie in der eingehenden Beschreibung dieser Erfindung erwähnt. Mittels herkömmlicher Züchtung ist eine Reihe von Mutanten-pflanzen entwickelt worden, die ein Spektrum an wünschenswerten Lipiden und Fettsäuren, Cofaktoren und Enzymen produzieren. Die Selektion neuer Pflanzensorten mit verbesserter Produktion
35 eines bestimmten Moleküls ist jedoch ein zeitaufwändiges und schwieriges Verfahren oder sogar unmöglich, wenn die Verbindung in der entsprechenden Pflanze nicht natürlich vorkommt, wie im Fall von mehrfach ungesättigten C₁₈-, C₂₀-Fettsäuren und C₂₂-Fett-säuren und solchen mit längeren Kohlenstoffketten.

Aufgrund der positiven Eigenschaften ungesättigter Fettsäuren hat es in der Vergangenheit nicht an Ansätzen gefehlt, Gene, die an der Synthese von Fettsäuren bzw. Triglyceriden beteiligt sind, für die Herstellung von Ölen in verschiedenen Organismen mit geändertem Gehalt an ungesättigten Fettsäuren verfügbar zu machen. So wird in WO 91/13972 und seinem US-Äquivalent eine Δ-9-Desaturase beschrieben. In WO 93/11245 wird eine Δ-15-Desaturase in WO 94/11516 wird eine Δ-12-Desaturase beansprucht. Δ-6-Desaturasen werden in WO 93/06712, US 5,614,393, 10 WO 96/21022 und WO 99/27111 beschrieben. Weitere Desaturasen werden beispielsweise in EP-A-0 550 162, WO 94/18337, WO 97/30582, WO 97/21340, WO 95/18222, EP-A-0 794 250, Stukey et al., J. Biol. Chem., 265, 1990: 20144-20149, Wada et al., Nature 347, 1990: 200-203 oder Huang et al., Lipids 34, 1999: 15 649-659 beschrieben. In WO 96/13591 wird eine Δ-6-Palmitoyl-ACP-Desaturase beschrieben und beansprucht. Die biochemische Charakterisierung der verschiedenen Desaturasen ist jedoch bisher nur unzureichend erfolgt, da die Enzyme als membrangebundene Proteine nur sehr schwer zu isolieren und charakterisieren sind 20 (McKeon et al., Methods in Enzymol. 71, 1981: 12141-12147, Wang et al., Plant Physiol. Biochem., 26, 1988: 777-792).

In Hefen konnte sowohl eine Verschiebung des Fettsäurespektrums zu ungesättigten Fettsäuren hin als auch eine Steigerung der 25 Produktivität nachgewiesen werden (siehe Huang et al., Lipids 34, 1999: 649-659, Napier et al., Biochem. J., Vol. 330, 1998: 611-614). Die Expression der verschiedenen Desaturasen in transgenen Pflanzen zeigte allerdings nicht den gewünschten Erfolg. Eine Verschiebung des Fettsäurespektrum zu ungesättigten Fett- 30 säuren hin konnte gezeigt werden, gleichzeitig zeigte sich aber, dass die Syntheseleistung der transgenen Pflanzen stark nachließ, das heißt gegenüber den Ausgangspflanzen konnten nur geringere Mengen an Ölen isoliert werden.

35 Weder in Hefen noch in Pflanzen werden natürlicherweise mehrfach ungesättigte C₂₀- und/oder C₂₂-Fettsäuren mit mindestens zwei Doppelbindungen im Fettsäuremolekül wie Arachidonsäure (ARA) und/oder Eicosapentaensäure (EPA) und/oder Docosahexaensäure (DHA) hergestellt.

40 Nach wie vor besteht daher ein großer Bedarf an neuen Genen, die für Enzyme kodieren, die an der Biosynthese ungesättigter Fettsäuren beteiligt sind und es ermöglichen, diese in einem technischen Maßstab herzustellen. Keines der bisher bekannten 45 biotechnologischen Verfahren zur Herstellung von mehrfach

ungesättigten Fettsäuren liefert die vorgenannten Fettsäuren in wirtschaftlich nutzbaren Mengen.

Bei der Expression von Genen in Pflanzen gibt es immer wieder Probleme, dass heißt es kommt durch die Expression nicht zur erwarteten Steigerung bei der Herstellung des gewünschten Wertprodukts.

Verschiedene Methoden zum Einschleusen von Genen in das Genom von Pflanzen sind bekannt (Halford NG, Shewry PR, Br. Med. Bull., 2000; 56: 62-73). Ziel ist die Herstellung von Pflanzen mit vorteilhaften, neuen Eigenschaften, zum Beispiel zur Steigerung der landwirtschaftlichen Produktivität, zur Qualitätssteigerung bei Nahrungsmitteln oder zur Produktion bestimmter Chemikalien oder Pharmazeutika (Dunwell JM, J. Exp. Bot., 2000: 51 Spec No: 487-96).

Eine Grundvoraussetzung für die transgene Expression bestimmter Gene ist die Bereitstellung pflanzenspezifischer Promotoren. Promotoren sind wichtige Werkzeuge in der Pflanzenbiotechnologie, um die Expression eines bestimmten Gens in einer transgenen Pflanze lokal und zeitlich zu steuern und so bestimmte Wesensmerkmale der Pflanze auszunutzen bzw. erst zu erzielen. Verschiedene Promotoren für diverse Pflanzenarten, bestimmte Pflanzengewebe und Entwicklungsstadien sind bekannt.

Verwendet werden zum Beispiel konstitutive Promotoren wie der Promotor der Nopalinsynthase aus Agrobacterium, der TR-Doppelpromotor oder der Promotor des 35S-Transkriptes des Blumenkohlmosaikvirus (CaMV) (Odell et al., Nature 1985: 313, 810-812), der OCS (Octopin Synthase) Promotor aus Agrobacterium, der Ubiquitin Promotor (Holtorf S et al., Plant Mol. Biol. 1995, 29:637-646), die Promotoren der vakuolärer ATPase Untereinheiten oder der Promotor eines prolinreichen Proteins aus Weizen (WO 91/13991). Nachteilig bei diesen Promotoren ist, dass sie in fast allen Geweben der Pflanze konstitutiv aktiv sind. Eine gezielte Expression von Genen in bestimmten Pflanzenteilen oder zu bestimmten Entwicklungszeitpunkten ist mit diesen Promotoren nicht möglich.

Promotoren, deren Aktivität gewebsspezifisch oder entwicklungsgabhängig reguliert ist, wurden isoliert. Spezifitäten sind für die Antheren, Ovarien, Blüten, Blätter, Stengel, Wurzeln und Samen beschrieben. Die Stringenz der Spezifität, als auch die Expressionsaktivität dieser Promotoren ist sehr unterschiedlich. Zu nennen sind Promotoren, die eine blattspezifische Expression gewährleisten, wie der Promotor der cytosolischen FBPase aus Kartoffel (WO 97/05900), der SSU Promotor (small subunit) der Rubi-

5

sco (Ribulose-1,5-bisphosphatcarboxylase) oder der ST-LSI Promotor aus Kartoffel (Stockhaus et al., EMBO J. 8 (1989), 2445 - 245).

5 Weitere Promotoren sind beispielsweise spezifische Promotoren für Knollen, Speicherwurzeln oder Wurzeln, wie beispielsweise der Patatin Promotor Klasse I (B33), der Promotor des Cathepsin D Inhibitors aus Kartoffel, der Promotor der Stärke Synthase (GBSS1) oder der Sporamin Promotor, fruchtspezifische Promotoren, wie
10 beispielsweise der fruchtspezifische Promotor aus Tomate (EP-A 409625), fruchtreifungspezifische Promotoren, wie beispielsweise der fruchtreifungspezifische Promotor aus Tomate (WO 94/21794), blütenspezifische Promotoren, wie beispielsweise der Phytoen-Synthase Promotor (WO 92/16635) oder der Promotor des P-rr Gens (WO
15 98/22593).

Eine Variation der Aktivität eines Promotoren anhängig vom Entwicklungsstadium der Pflanze wird unter anderem von Baerson et al. beschrieben (Baerson SR, Lamppa GK. Plant Mol Biol.
20 1993;22(2):255-67).

Samenspezifische Promotoren sind aufgrund der Bedeutung des Samens als eine der Hauptnahrungs- bzw. Futterquellen von Mensch und Tier und als Produktionsort für Wertstoffe von besonderen Interesse. Bekannt sind Promotoren, die eine Expression in Samen und pflanzlichen Embryonen steuern. So wurden beispielsweise die Promotoren von Genen identifiziert, die für Speicherproteine verschiedener Dicotyledonen kodieren. Samenspezifische Promotoren sind zum Beispiel der Promotor des Phaseolins (US 5504200, Bustos
25 MM et al., Plant Cell. 1989;1(9):839-53), des 2S Albumingens (Josephson LG et al., J Biol Chem 1987, 262:12196-12201), des Legumins (Shirsat A et al., Mol Gen Genet. 1989;215(2):326-331), des USP (unknown seed protein; Bäumlein H et al., Molecular & General Genetics 1991, 225(3):459-67) des Napin Gens (Stalberg K, et
30 al., L. Planta 1996, 199:515-519), des Saccharosebindeproteins (WO 00/26388) oder der LeB4-Promotor (Bäumlein H et al., Mol Gen genet 1991, 225: 121-128). Diese steuern eine hohe samenspezifische Expression von Speicherproteinen.

40 Trotz der allgemeinen Annahme, daß pflanzliche Promotoren von einer Spezies auf die andere Übertragbar sind und auch in artfremden Pflanzenspezies ähnliche Aktivitäten und Spezifitäten aufweisen, mehren sich Hinweise auf Einschränkungen von dieser Annahme. So zeigte es sich, daß die Höhe der transgenen Expression von heterologen Genen unter Kontrolle dieser Promotoren, oft stark von der Art der Wirtspflanze abhängig ist. Es wurde festgestellt, dass die Expression nicht immer absolut zelltypspezifisch ist.

Unterschiede im Expressionsmuster und der Expressionsstärke eines bestimmten Promotors können durch unterschiedliche Wirtspflanzen oder durch unterschiedliche Insertionsorte in das Genom der Wirtspflanze bedingt sein (Goossens A et al., Plant Phys 1999, 5 120:1095-1104).

Durch eine Genexpression in anderen Pflanzenteilen kann die Verwendung eines Promoters in einer anderen Pflanzenart sehr eingeschränkt sein. Etwa wenn die Expression des Genes in den Metabolismus der Zelle, die Zusammensetzung der Membranlipide oder die Biosynthese eingreift.

Es bestand daher die Aufgabe weitere Expressionskassetten für die Expression in Pflanzen zur Verfügung zu stellen. Und diese für die Expression von Genen vorteilhaft Biosynthesegenen allein oder gegebenenfalls in Kombination mit anderen Enzymen in einem Verfahren beispielsweise zur Herstellung mehrfach ungesättigter Fettsäuren zu verwenden. Diese Aufgabe wurde durch die erfundungsgemäßen Expressionskassette mit einer Struktur ausgewählt aus der Gruppe gelöst:

- a) L1 - Promotor - Strukturgen - L2,
- b) L1 - Promotor - Strukturgen - L2 - L1 - Promotor - Strukturgen - L2,
- c) L1 - Promotor - Strukturgen - L2 - L1 - Promotor - Strukturgen - L2 - L1 - Promotor - Strukturgen - L2,

wobei L1, L2, Promotor und Strukturgen die folgende Bedeutung hat:

L1 = SEQ ID NO: 32 oder eine äquivalente Restriktionsschnittstellen enthaltende Sequenz,

35

L2 = unabhängig voneinander SEQ ID NO: 33, 34 oder 35 oder äquivalente Restriktionsschnittstellen enthaltende Sequenzen,

Promotor = pflanzlicher Promotor

40

Strukturgen = eine in Pflanzen exprimierbare Nukleinsäuresequenz.

Vorteilhaft ist das Strukturgen ein Biosynthesegen, bevorzugt ist es ein Biosynthesegen des Lipid- oder Fettsäurestoffwechsels, vorteilhaft ein pflanzliches Gen. In einer besonders vorteilhaften Ausführungsform ist das Strukturgen eine Nukleinsäuresequenz,

die für Proteine ausgewählt aus der Gruppe :

Acyl-CoA-Dehydrogenase(n), Acyl-ACP[= acyl carrier protein]-Desaturase(n), Acyl-ACP-Thioesterase(n), Fettsäure-Acyl-Transferase(n), Fettsäure-Synthase(n), Fettsäure-Hydroxylase(n), Acetyl-

5 Coenzym A-Carboxylase(n), Acyl-Coenzym A-Oxidase(n), Fettsäure-Desaturase(n), Fettsäure-Acetylensäuren, Lipoxygenasen, Triacylglycerol-Lipasen, Allenoxid-Synthasen, Hydroperoxid-Lyasen oder Fettsäure-Elongase(n), kodiert.

10 Ganz besonders bevorzugt ist das Strukturen eine Nukleinsäuresequenz ausgewählt aus der Gruppe:

a) einer Nukleinsäuresequenz mit der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 11 dargestellten Sequenz,

15 b) Nukleinsäuresequenzen, die aufgrund des degenerierten genetischen Codes durch Rückübersetzung der in SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6 oder SEQ ID NO: 12 dargestellten Aminosäuresequenzen erhalten werden,

20 c) Derivate der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 11 dargestellten Nukleinsäuresequenz, die für Polypeptide mit der in SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6 oder SEQ ID NO: 12 dargestellten Aminosäuresequenzen codieren
25 und mindestens 50 % Homologie auf Aminosäureebene aufweisen, ohne daß die enzymatische Wirkung der Polypeptide wesentlich reduziert ist.

Unter einer äquivalenten Restriktionsschnittstellen enthaltende
30 Sequenz im Sinne der Erfindung sind Sequenzen zu verstehen, die Restriktionsschnittstellen enthalten, die für den Aufbau multiplexter Expressionskassetten geeignet sind, das heißt geeigneter Weise nicht im Strukturen vorhanden sind oder im binären Vektor. Solche Restriktionsschnittstellen wie beispielhaft EcoRI, BamHI,
35 SacI, PstI, NcoI, NdeI, BglI, BglII, XbaI, Xba und weitere sind dem Fachmann bekannt und können einschlägigen Fachbüchern entnommen werden.

Die erfindungsgemäßen Expressionskassetten können für die Expression von Genen in wirtschaftlich wichtigen Kulturpflanzen wie beispielsweise Lein, für die keine endogenen samenspezifischen Promotoren bekannt waren, verwendet werden. Lein ist, wie die vorliegenden Arbeiten zeigten, für eine samenspezifische Expression von Genen besonders problematisch, da offensichtlich mehrere

40 Promotoren, die vom Fachmann für samenspezifische Expression in anderen Pflanzen routinemäßig benutzt werden, in z.B. in anderen Pflanzen wie Lein nicht funktionieren, das heißt nicht zu einer

Transkription bzw. letztlich zu einer Expression der mRNA des Strukturgens führt.

Verwendung der oben genannten erfindungsgemäßen Expressionskas-
5 setten in einem Verfahren zur Herstellung von Fettsäureestern mit einem erhöhten Gehalt an mehrfach ungesättigten Fettsäuren mit mindestens zwei Doppelbindungen dadurch gekennzeichnet, daß man mindestens eine Nukleinsäuresequenz ausgewählt aus der Gruppe

10 a) einer Nukleinsäuresequenz mit der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 11 dargestellten Sequenz,

15 b) Nukleinsäuresequenzen, die aufgrund des degenerierten genetischen Codes durch Rückübersetzung der in SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6 oder SEQ ID NO: 12 dargestellten Aminosäuresequenzen erhalten werden,

20 c) Derivate der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 11 dargestellten Nukleinsäuresequenz, die für Polypeptide mit der in SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6 oder SEQ ID NO: 12 dargestellten Aminosäuresequenzen codieren und mindestens 50 % Homologie auf Aminosäureebene aufweisen, ohne dass die enzymatische Wirkung
25 der Polypeptide wesentlich reduziert ist,

in einen Fettsäureester produzierenden Organismus einbringt, anzieht und die dem Organismus enthaltenden Fettsäureester isoliert.

30 Bei den im Verfahren verwendeten Nukleinsäuresequenzen handelt es sich um isolierte Nukleinsäuresequenzen, die für Polypeptide mit Δ-5-, Δ-6- oder Δ-12-Desaturaseaktivität codieren.

35 Im Verfahren werden vorteilhaft Fettsäureester mit mehrfach ungesättigten C₁₈-, C₂₀- und/oder C₂₂-Fettsäuremolekülen mit mindestens zwei Doppelbindungen im Fettsäureester hergestellt. Bevorzugt enthalten diese Fettsäuremoleküle drei, vier oder fünf Doppelbindungen und führen vorteilhaft zur Synthese von Arachidonsäure (ARA), Eicosapentaensäure (EPA) oder Docosahexaensäure (DHA).

Die Fettsäureester mit mehrfach ungesättigten C₁₈-, C₂₀- und/oder C₂₂-Fettsäuremolekülen können aus den Organismen, die für die Herstellung der Fettsäureester verwendet wurden, in Form eines Öls oder Lipids beispielsweise in Form von Verbindungen wie Sphingolipide, Phosphoglyceride, Lipide, Glycolipide, Phospholipide,

Monoacylglyceride, Diacylglyceride, Triacylglyceride oder sonstige Fettsäureester, die die mehrfach ungesättigten Fett- säuren mit mindestens zwei Doppelbindungen enthalten, isoliert werden.

5

Als Organismus für die Herstellung im Verfahren kommen prinzipiell alle prokaryontischen oder eukaryontischen Organismen wie prokaryontische oder eukaryontische Mikroorganismen wie gram-positive oder gram-negative Bakterien, Pilze, Hefen, Algen, Ciliaten, tierische oder pflanzliche Zellen, Tiere oder Pflanzen wie Moose, zweikeimblättrige oder einkeimblättrige Pflanzen in Frage. Vorteilhaft werden Organismen im erfindungsgemäßen Verfahren verwendet, die zu den Ölproduzierenden Organismen gehören, das heißt die für die Herstellung von Ölen verwendet werden, wie Mikroorganismen wie Cryptecodinium, Thraustochytrium, Phaeodactylum und Mortierella, Entomophthora, Mucor, Cryptecodinium sowie andere Algen oder Pilze sowie Tiere oder Pflanzen, insbesondere Pflanzen bevorzugt Ölfruchtpflanzen, die große Mengen an Lipidverbindungen enthalten, wie Sojabohne, Erdnuss, Raps, Canola, Sonnenblume, Safflor, Nachtkerze, Lein, Soja, Borretsch, Bäume (Ölpalme, Kokosnuss) oder Feldfrüchte, wie Mais, Weizen, Roggen, Hafer, Triticale, Reis, Gerste, Baumwolle, Maniok, Pfeffer, Tagetes, Solanaceen-Pflanzen, wie Kartoffel, Tabak, Aubergine und Tomate, Vicia-Arten, Erbse, Alfalfa oder Buschpflanzen (Kaffee, Kakao, Tee), Salix-Arten, Bäume (Ölplame, Kokosnuß) sowie ausdauernde Gräser und Futterfeldfrüchte. Besonders bevorzugte erfindungsgemäße Pflanzen sind Ölfruchtpflanzen, wie Soja, Erdnuß, Raps, Canola, Lein, Nachtkerze, Sonnenblume, Safflor oder Bäume (Ölpalme, Kokosnuß).

30

Das Verfahren beinhaltet entweder die Züchtung eines geeigneten transgenen Organismus bzw. transgenen Mikroorganismus oder die Züchtung von transgenen Pflanzenzellen, -geweben, -organen oder ganzen Pflanzen, umfassend die erfindungsgemäßen Nukleotidsequenzen der SEQ ID NO: 1, 3, 5 oder 11 gegebenenfalls in Verbindungen mit den in SEQ ID NO: 7 und/oder SEQ ID NO: 9 dargestellten Sequenzen allein oder in Kombination mit Sequenzen von vorteilhaften erfindungsgemäßen Expressionskassetten in vorteilhaften Vektoren mit SEQ ID NO: 13-17 oder ihre Homologen, Derivate oder Analoga oder ein Genkonstrukt, das die SEQ ID NO: 1, 3, 5 oder 11 ggf. in Verbindung mit SEQ ID NO: 7 und/oder 9 oder ihre Homologen, Derivate oder Analoga umfasst, oder einen Vektor, der diese Sequenz oder das Genkonstrukt umfasst, welches die Expression erfindungsgemäßer Nukleinsäuremoleküle herbeiführt, so dass eine Feinchemikalie produziert wird. Bei einer bevorzugten Ausführungsform umfasst das Verfahren ferner den Schritt des Gewinnens einer Zelle, die eine solche erfindungsgemäße Nukleinsäuresequen-

10

zen enthält, wobei eine Zelle mit einer Desaturasenukleinsäuresequenz, einem Genkonstrukt oder einem Vektor, welche die Expression einer erfindungsgemäßen Desaturasenukleinsäure allein oder in Kombination herbeiführen, transformiert wird. Bei einer 5 weiteren bevorzugten Ausführungsform umfasst dieses Verfahren ferner den Schritt des Gewinnens der Feinchemikalie aus der Kultur. Bei einer besonders bevorzugten Ausführungsform gehört die Zelle zur Ordnung der Ciliaten, zu Mikroorganismen, wie Pilzen, oder zum Pflanzenreich, insbesondere zu Ölfruchtpflanzen, 10 besonders bevorzugt sind Mikroorganismen oder Ölfruchtpflanzen beispielsweise Erdnuss, Raps, Canola, Lein, Soja, Safflower (Distel), Sonnenblumen oder Borretsch.

Unter transgen im Sinne der Erfindung ist zu verstehen, daß die 15 im Verfahren verwendeten Nukleinsäuren oder die erfindungsgemäßen Expressionskassetten nicht an ihrer natürlichen Stelle im Genom eines Organismus sind, dabei können die Nukleinsäuren homolog oder heterolog exprimiert werden. Transgen bedeutet aber auch, dass die Nukleinsäuren oder Expressionskassetten an ihrem natürlichen Platz im Genom eines Organismus sind, dass jedoch die Sequenz gegenüber der natürlichen Sequenz verändert wurde und/oder das die Regulationssequenzen, der natürlichen Sequenzen verändert wurden. Bevorzugt ist unter transgen die Expression der erfindungsgemäßen Nukleinsäuren an nicht natürlicher Stelle im Genom 20 zu verstehen, das heißt eine homologe oder bevorzugt heterologe Expression der Nukleinsäuren liegt vor. Bevorzugte transgene Organismen sind die oben genannten transgenen Pflanzen bevorzugt Ölfruchtpflanzen.

30 Aus den im Verfahren hergestellten Fettsäureestern lassen sich die enthaltenden mehrfach ungesättigten Fettsäuren beispielsweise über eine Alkalibehandlung wie wässrige KOH oder NaOH vorteilhaft in Gegenwart eines Alkohols wie Methanol oder Ethanol freisetzen und isolieren über beispielsweise Phasentrennung und anschließen- 35 der Ansäuerung über z.B. H₂SO₄.

Die im Verfahren hergestellten Fettsäureester fallen in Form von Ölen, Lipiden und/oder Fettsäuren an, die mindestens zwei Doppelbindungen in den Fettsäremolekülen bevorzugt drei, vier, fünf 40 oder sechs Doppelbindungen enthalten. Auch sind Zusammensetzungen, die die genannten Öl-, Lipid- und/oder Fettsäuren enthalten, sowie die Verwendung der Öle, Lipide und/oder Fettsäuren oder der Zusammensetzungen in Futter, Nahrungsmitteln, Kosmetika oder Pharmazeutika ein weiterer Anwendungsmöglichkeit der vorgenannten 45 Stoffe.

Ein weiterer Aspekt betrifft ein Verfahren zur Modulation der Produktion eines Moleküls durch einen Mikroorganismus. Diese Verfahren umfassen das Zusammenbringen der Zelle mit einer Substanz, welche die erfindungsgemäßen Desaturaseaktivität allein oder in 5 Kombination oder die Desaturasenukleinsäureexpression moduliert, so dass eine zellassoziierte Aktivität relativ zu der gleichen Aktivität in Abwesenheit der Substanz verändert wird. Bei einer bevorzugten Ausführungsform wird/werden ein oder zwei Stoffwechselweg(e) der Zelle für Lipide und Fettsäuren, Cofaktoren und Enzyme moduliert oder der Transport von Verbindungen über diese Membranen moduliert, so dass die Ausbeute oder die Rate der Produktion einer gewünschten Feinchemikalie durch diesen Mikroorganismus verbessert ist. Die Substanz, welche die Desaturaseaktivität moduliert, kann eine Substanz sein, welche die 10 Desaturaseaktivität oder Desaturasenukleinsäureexpression stimuliert oder die als Zwischenprodukt bei der Fettsäurebiosynthese verwendet werden kann. Beispiele für Substanzen, welche die Desaturaseaktivität oder Desaturasenukleinsäureexpression stimulieren, sind u.a. kleine Moleküle, aktive Desaturasen sowie 15 desaturasencodierende Nukleinsäuren, die in die Zelle eingebracht worden sind. Beispiele für Substanzen, welche die Desaturaseaktivität oder -Expression hemmen, sind u.a. kleine Moleküle und Antisense- Desaturasenukleinsäuremoleküle.

20 Ein weiterer Aspekt betrifft ein Verfahren zur Modulation der Ausbeuten einer gewünschten Verbindung aus einer Zelle, umfassend das Einbringen eines Wildtyp- oder Mutanten-Desaturasegens, das entweder auf einem separaten Plasmid gehalten oder in das Genom der Wirtszelle integriert wird, in eine Zelle. Bei Integration in 25 das Genom kann die Integration zufallsgemäß sein oder durch derartige Rekombination erfolgen, dass das native Gen durch die eingebrachte Kopie ersetzt wird, wodurch die Produktion der gewünschten Verbindung durch die Zelle moduliert wird, oder durch Verwendung eines Gens in trans, so dass das Gen mit einer funktionellen Expressionseinheit, welche mindestens eine die Expression eines Gens gewährleistende Sequenz und mindestens eine die Polyadenylierung eines funktionell transkribierten Gens gewährleistende Sequenz enthält, funktionell verbunden ist.

30 Bei einer bevorzugten Form des Verfahrens sind die Ausbeuten modifiziert. Bei einer weiteren bevorzugten Ausführungsform wird die gewünschte Chemikalie vermehrt, wobei unerwünschte störende Verbindungen vermindert werden können. Bei einer besonders bevorzugten Ausführungsform ist die gewünschte Feinchemikalie ein Lipid oder eine Fettsäure, ein Cofaktor oder ein Enzym. Bei besonders bevorzugten Ausführungsform ist diese 35 Chemikalie eine mehrfach ungesättigte Fettsäure. Stärker bevor-

12

zugt ist sie ausgewählt aus Arachidonsäure (ARA), Eicosapentaensäure (EPA) oder Docosahexaensäure (DHA).

Die vorliegende Erfindung stellt vorteilhafte Multiexpressionskassetten und Konstrukte zur multiparallelen samenspezifischen Expression von Genkombinationen in Pflanzen zur Verfügung.

Diese können im oben beschriebenen Verfahren zur Expression von Genen, bevorzugt den in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, 10 SEQ ID NO: 7, SEQ ID NO: 9 oder SEQ ID NO: 11 beschriebenen, in Algen und Pilzen und Pflanzen, insbesondere Ölfruchtpflanzen sind bevorzugte Organismen für das Verfahren verwendet werden.

Durch die Verwendung der erfindungsgemäßen Expressionskassetten 15 können im Verfahren in Verbindung mit den oben genannten Nukleinsäuremoleküle zur gentechnologisch Veränderung von Pflanzen verwendet werden, so dass sie schließlich zur Herstellung von besseren oder effizienteren Produzenten einer oder mehrerer Feinchemikalien führen. Diese verbesserte Produktion oder Effizienz der 20 Produktion einer Feinchemikalie kann durch eine direkte Wirkung der Manipulation eines erfindungsgemäßen Gens oder durch eine indirekte Wirkung dieser Manipulation hervorgerufen werden. Unter Feinchemikalien sind im Sinne der Erfindung beispielsweise Fettsäureester, die mehrfach ungesättigte Fettsäuren mit mindestens 25 zwei Doppelbindungen enthalten wie Sphingolipide, Phosphoglyceride, Lipide, Glycolipide, Phospholipide, Monoacylglyceride, Diacylglyceride, Triacylglyceride oder sonstige Fettsäureester, die die mehrfach ungesättigten Fettsäuren mit mindestens zwei Doppelbindungen enthalten, zu verstehen. Weiter sind darunter zu verstehen Verbindungen wie Vitamine beispielsweise Vitamin E, Vitamin C, Vitamin B2, Vitamin B6, Pantolacton, Carotinoide wie Astaxanthin, β-Carotin, Zeaxanthin und andere.

Moose und Algen sind die einzigen bekannten Pflanzensysteme, 35 die erhebliche Mengen an mehrfach ungesättigten Fettsäuren, wie Arachidonsäure (ARA) und/oder Eicosapentaensäure (EPA) und/oder Docosahexaensäure (DHA) herstellen. Moose enthalten PUFAs in Membranlipiden während Algen, algenverwandte Organismen und einige Pilze auch nennenswerte Mengen an PUFAs in der Triacyl-glycerolfraktion akkumulieren. Daher eignen sich Nukleinsäuremoleküle, die aus solchen Stämmen isoliert werden, die PUFAs auch 40 in der Triacylglycerolfraktion akkumulieren, besonders vorteilhaft zur Modifikation des Lipid- und PUFA-Produktionssystems in einem Wirt, insbesondere in Mikroorganismen, wie den vorstehend erwähnten Mikroorganismen, und Pflanzen, wie Ölfruchtpflanzen,

beispielsweise Raps, Canola, Lein, Soja, Sonnenblumen, Borretsch. Sie sind deshalb vorteilhaft im Verfahren verwendbar.

Die im Verfahren unter Verwendung der erfindungsgemäßen Expressionskassetten verwendeten Nukleinsäuresequenzen codieren für Desaturasen, die zur Produktion langkettiger mehrfach ungesättigter Fettsäuren, vorzugsweise mit mehr als sechzehn, achtzehn oder zwanzig Kohlenstoffatomen im Kohlenstoffgrundgerüst der Fettsäure und/oder mindestens zwei Doppelbindungen in der Kohlenstoffkette, geeignet sind, wobei eine Nukleinsäure für ein Enzym codiert, das Doppelbindungen in die Δ -5-Position, in einem anderen Fall in die Δ -6-Position und in einem weiteren Fall in die Δ -12-Position einführen kann. Mithilfe dieser Nukleinsäuren können hohe Mengen an PUFAs in der Triacylglycerolfraktion erhalten werden. Weiterhin wurden weitere Desaturasen isoliert, die allein oder zusammen mit einer Δ -4-Desaturase für ein Verfahren zur Produktion polyungesättigter Fettsäuren genutzt werden können. Dabei ist in der Anmeldung unter dem Singular d.h. unter einem Desaturasegen oder -Protein auch der Plural d.h. die Desaturasegenen oder -Proteinen zu verstehen.

Die Herstellung einer Triensäure mit C₁₈-Kohlenstoffkette mithilfe von Desaturasen konnte bisher gezeigt werden. In diesen literaturbekannten Verfahren wurde die Herstellung von γ -Linolensäure beansprucht. Bisher konnte jedoch niemand die Herstellung sehr langkettiger mehrfach ungesättigter Fettsäuren (mit C₂₀- und längerer Kohlenstoffkette sowie von Triensäuren und höher ungesättigten Typen) allein durch modifizierte Organismen zeigen.

Zur Herstellung der erfindungsgemäßen langkettiger PUFAs müssen die mehrfach ungesättigten C₁₈-Fettsäuren zunächst durch die enzymatische Aktivität einer Elongase um mindestens zwei Kohlenstoffatome verlängert werden. Nach einer Elongationsrunde führt diese Enzymaktivität zu C₂₀-Fettsäuren, und nach zwei, drei und vier Elongationsrunden zu C₂₂-, C₂₄- oder C₂₆-Fettsäuren. Die in dieser Erfindung offenbarten Nukleinsäuresequenzen, die für verschiedene Desaturasen codieren, können im Konzert mit Elongasen zu sehr langkettigen, polyungesättigten führen. Die Aktivität der erfindungsgemäßen Desaturasen führt vorzugsweise zu C₁₈-, C₂₀- und/oder C₂₂-Fettsäuren mit mindestens zwei Doppelbindungen im Fettsäuremolekül, vorzugsweise mit drei, vier, fünf oder sechs Doppelbindungen, besonders bevorzugt zu C₁₈- und/oder C₂₀-Fettsäuren mit mindestens zwei Doppelbindungen im Fettsäuremolekül, vorzugsweise mit drei, vier oder fünf Doppelbindungen im Molekül. Die Fettsäureelongation kann durch Kombination der erfindungsgemäßen Desaturasen mit einer Elongaseaktivität erfolgen, wobei die durch die in SEQ ID NO: 9 codierte Elongase vorteilhaft verwendet

werden kann. Nachdem die Verlängerung mit dem erfindungsgemäßen Enzym(en) stattgefunden hat, können weitere Desaturierungsschritte wie z.B. eine solche in Δ -5-Position erfolgen. Auch die Kombination mit anderen Elongasen wie solche, die zu einer Verlängerung von C₁₈- auf C₂₀- oder von C₂₀- auf C₂₂₋₂₄ Ketten wie in WO0012720 offenbart führt, kann Verwendung finden und/oder einer Desaturase mit Aktivität für Δ -4-Position kann vorteilhaft eingesetzt werden, um die hoch desaturierten Fettsäuren zu erhalten. Daher führen die Produkte der Desaturaseaktivitäten und der möglichen weiteren Desaturierung zu bevorzugten PUFAs mit einem höheren Desaturierungsgrad, wie Dihomo-gamma-Lonolensäure, Docosadiensäure, Arachidonsäure, ω 6-Eicosatrienidihomo- γ -linolensäure, Eicosapentaensäure, ω 3-Eicosatriensäure, ω 3-Eicosatetraensäure, Docosapentaensäure oder Docosahexaensäure. Substrate der erfindungsgemäßen Enzymaktivität sind zum Beispiel Taxolsäure; 6,9-Octadecadiensäure, Linolsäure, Pinolensäure, α -oder γ -Linolensäure oder Stearidonsäure sowie Arachidonsäure, Eicosatetraensäure, Docosapentaensäure, Eicosapentaensäure. Bevorzugte Substrate sind Linolsäure, γ -Linolensäure und/oder α -Linolensäure sowie Arachidonsäure, Eicosatetraensäure, Docosapentaensäure, Eicosapentaensäure. Besonders bevorzugt als Produkte des erfindungsgemäßen Verfahrens sind Arachidonsäure, Docosapentaensäure, Eicosapentaensäure. Die C₁₈-Fettsäuren mit mindestens zwei Doppelbindungen in der Fettsäure können durch die erfindungsgemäße enzymatische Aktivität in Form der freien Fettsäure oder in Form der Ester, wie Phospholipide, Glycolipide, Sphingolipide, Phosphoglyceride, Monoacylglyceride, Diacylglyceride oder Triacylglyceride, verlängert werden.

Für die menschliche Ernährung ist konjugierte Linolsäure "CLA" von besonderer Bedeutung. Unter CLA versteht man insbesondere Fettsäuren wie C_{18:2} ^{9 cis, 11trans} oder das Isomer C_{18:2} ^{10trans, 12 cis}, die aufgrund menschlicher Enzymsysteme nach Aufnahme im Körper desaturiert bzw. elongiert werden können und zu gesundheitsfördernden Effekten beitragen können. Mit den erfindungsgemäßen Desaturasen (Δ -12-Desaturase) können auch solche konjugierten Fettsäuren mit wenigstens zwei Doppelbindungen im Molekül desaturiert werden und damit solche gesundheitsfördernden Fettsäuren der menschlichen Ernährung zugeführt werden. Weitere Beispiel für konjugierte Fettsäuren sind alpha-Parinarsäure, Punicasäure, Eleostearinsäure und Calendulasäure.

Unter der Verwendung von der erfindungsgemäßen Expressionskassetten in Klonierungsvektoren in Pflanzen und bei der Pflanzentransformation, wie denjenigen, die veröffentlicht sind in und dort zitiert sind: Plant Molecular Biology and Biotechnology (CRC Press, Boca Raton, Florida), Kapitel 6/7, S. 71-119 (1993); F.F.

White, Vectors for Gene Transfer in Higher Plants; in: Transgenic Plants, Bd. 1, Engineering and Utilization, Hrsgb.: Kung und R. Wu, Academic Press, 1993, 15-38; B. Jenes et al., Techniques for Gene Transfer, in: Transgenic Plants, Bd. 1, Engineering and Utilization, Hrsgb.: Kung und R. Wu, Academic Press (1993), 128-143; Potrykus, Annu. Rev. Plant Physiol. Plant Molec. Biol. 42 (1991), 205-225)), lassen sich Gene vorteilhaft Biosynthesegene wie die oben beschriebenen Nukleinsäuren zur gentechnologischen Veränderung eines breiten Spektrums an Pflanzen verwenden, so dass 5 diese ein besserer oder effizienterer Produzent beispielsweise eines oder mehrerer von Lipiden hergeleiteter Produkte, wie PU-FAs, werden. Diese verbesserte Produktion oder Effizienz der Produktion eines beispielsweise von Lipiden hergeleiteten Produktes, wie PUFA, kann durch direkte Wirkung der Manipulation oder eine 10 indirekte Wirkung dieser Manipulation hervorgerufen werden.

15

Es gibt eine Reihe von Mechanismen, durch die die Veränderung eines erfindungsgemäßen Desaturaseproteins die Ausbeute, Produktion und/oder Effizienz der Produktion einer Fein- 20 chemikalie aus einer Ölfruchtpflanze oder einem Mikroorganismus aufgrund eines veränderten Proteins direkt beeinflussen kann. Die Anzahl oder Aktivität des Desaturaseproteins oder -Gens sowie von Genkombinationen von Desaturasen und Elongasen kann erhöht sein, so dass größere Mengen dieser Verbindungen de 25 novo hergestellt werden, weil den Organismen diese Aktivität und Fähigkeit zur Biosynthese vor dem Einbringen des entsprechenden Gens fehlte. Entsprechendes gilt für die Kombination mit weiteren Desaturasen oder Elongasen oder weiteren Enzymen aus dem Lipidstoffwechsel. Auch die Verwendung verschiedener 30 divergenter, d.h. auf DNA-Sequenzebene unterschiedlicher Sequenzen kann dabei vorteilhaft sein bzw. die Verwendung von Promotoren zur Genexpression, die eine andere zeitliche Genexpression z.B. abhängig vom Reifegrad eines Samens oder Öl-speichernden Gewebes ermöglicht.

35

Durch das Einbringen eines Desaturasegens oder mehrerer Desaturasegene unter Kontrolle der erfindungsgemäßen Expressionskassetten in einen Organismus allein oder in Kombination mit anderen Genen in eine Zelle kann nicht nur den Biosynthesefluss zum Endprodukt 40 erhöht, sondern auch die entsprechende Zusammensetzung der Endprodukte beispielsweise der Triacylglycerine erhöht oder de novo geschaffen werden. Ebenso kann die Anzahl oder Aktivität anderer Gene, die am Import von Nährstoffen, die zur Biosynthese einer oder mehrerer Feinchemikalien (z.B. Fettsäuren, polaren und neutralen Lipiden) nötig sind, erhöht sein, so dass die Konzentration dieser Vorläufer, Cofaktoren oder Zwischenverbindungen innerhalb der Zellen oder innerhalb des Speicherkompartiments er-

höht ist, wodurch die Fähigkeit der Zellen zur Produktion von PU-FAs, wie im folgenden beschrieben, weiter gesteigert wird. Fettsäuren und Lipide sind selbst als Feinchemikalien wünschenswert; durch Optimierung der Aktivität oder Erhöhung der Anzahl einer 5 oder mehrerer Desaturasen, die an der Biosynthese dieser Verbindungen beteiligt sind, oder durch Zerstören der Aktivität einer oder mehrerer Desaturasen, die am Abbau dieser Verbindungen beteiligt sind, kann es möglich sein, die Ausbeute, Produktion und/ oder Effizienz der Produktion von Fettsäure- und Lipidmolekülen 10 aus Pflanzen oder Mikroorganismen zu steigern.

Die Mutagenese der/des erfindungsgemäßen Desaturasegene(s) kann weiterhin zu einem Desaturaseprotein mit geänderten Aktivitäten führen, welche die Produktion einer oder mehrerer gewünschter 15 Feinchemikalien direkt oder indirekt beeinflussen. Beispielsweise kann die Anzahl oder Aktivität der/des erfindungsgemäßen Desaturasegens(e) gesteigert werden, so dass die normalen Stoffwechselabfälle oder -nebenprodukte der Zelle (deren Menge möglicherweise aufgrund der Überproduktion der gewünschten Fein- 20 chemikalie erhöht ist) effizient exportiert werden, bevor sie andere Moleküle oder Prozesse innerhalb der Zelle (welche die Lebensfähigkeit der Zelle senken würden) zerstören oder die Biosynthesewege der Feinchemikalie stören würden (wodurch die Ausbeute, Produktion oder Effizienz der Produktion der gewünschten 25 Feinchemikalie verringert wird). Ferner können die relativ großen intrazellulären Mengen der gewünschten Feinchemikalie selbst toxisch für die Zelle sein oder Enzym-Rückkopplungsmechanismen, wie die allosterische Regulation, stören, beispielsweise könnte sie durch Steigerung der Aktivität oder Anzahl anderer strom- 30 abwärts folgender Enzyme oder Entgiftungsenzyme des PUFA-Wegs die Allokation der PUFA in die Triacylglycerin-Fraktion steigern, man könnte die Lebensfähigkeit von Saatzellen erhöhen, was wiederum zu besserer Entwicklung von Zellen in Kultur oder zu Saaten führt, die die gewünschte Feinchemikalie produzieren.

35 Das erfindungsgemäße Desaturasegen kann auch so manipuliert werden, dass die entsprechenden Mengen der verschiedenen Lipid- und Fettsäuremoleküle hergestellt werden. Dies kann eine einschneidende Wirkung auf die Lipidzusammensetzung der Membran der Zelle haben und erzeugt neue Öle zusätzlich zum Auftreten 40 neusynthetisierter PUFAs. Da jeder Lipidtyp unterschiedliche physikalische Eigenschaften hat, kann eine Veränderung der Lipidzusammensetzung einer Membran die Membranfluidität erheblich verändern. Änderungen der Membranfluidität können sich auf den Transport von Molekülen über die Membran sowie auf die 45 Unversehrtheit der Zelle auswirken, die beide eine entscheidende Wirkung auf die Produktion von Feinchemikalien besitzen. In Pflanzen können diese Änderungen überdies auch andere Merk-

male, wie Toleranz gegenüber abiotischen und biotischen Stress-situationen, beeinflussen.

Im Verfahren können isolierte Nukleinsäuremoleküle (z.B. cDNAs),
5 umfassend Nukleotidsequenzen, die eine Desaturase oder mehrere
Desaturasen oder biologisch aktive Teile davon codieren, oder Nu-
kleinsäurefragmente, die sich als Primer oder Hybridisierungsson-
den zum Nachweis oder zur Amplifikation desaturasekodierender
Nukleinsäuren (z.B. DNA oder mRNA) eignen, verwendet werden. Bei
10 besonders bevorzugten Ausführungsformen umfasst das Nukleinsäure-
molekül eine der in Sequenz ID NO: 1 bzw 3 und 5 dargestellten Nu-
kleotidsequenzen oder die kodierende Region oder ein Komplement
einer dieser Nukleotidsequenzen. Bei anderen besonders bevorzug-
ten Ausführungsformen umfasst das isolierte Nukleinsäuremolekül
15 eine Nukleotidsequenz, die an eine Nukleotidsequenz, wie in der
Sequenz SEQ ID NO: 1, 3, 5 oder 11 dargestellt, oder einen Teil
davon hybridisiert oder zu mindestens etwa 50 %, vorzugsweise
mindestens etwa 60 %, stärker bevorzugt mindestens etwa 70 %,
80 % oder 90 % und noch stärker bevorzugt mindestens etwa 95 %,
20 96 %, 97 %, 98 %, 99 % oder mehr homolog dazu ist. Bei anderen
bevorzugten Ausführungsformen kodiert das isolierte Nukleinsäure-
molekül eine der in der Sequenz SEQ ID NO: 2, 4, 6 oder 12 darge-
stellten Aminosäuresequenzen. Das bevorzugte Desaturasegen be-
sitzt vorzugsweise auch mindestens eine der hier beschriebenen
25 Desaturaseaktivitäten.

Bei einer weiteren Ausführungsform kodiert das isolierte
Nukleinsäuremolekül ein Protein oder einen Teil davon, wobei
das Protein oder der Teil davon eine Aminosäuresequenz enthält,
30 die ausreichend homolog zu einer Aminosäuresequenz der Sequenz
SEQ ID NO: 2, 4, 6 oder 12 ist, dass das Protein oder der Teil
davon eine Desaturaseaktivität beibehält. Vorzugsweise behält das
Protein oder der Teil davon, das/der von dem Nukleinsäuremolekül
kodiert wird, die Fähigkeit, am Stoffwechsel von zum Aufbau
35 von Zellmembranen von Pflanzen notwendigen Verbindungen oder
am Transport von Molekülen über diese Membranen teilzunehmen.
Bei einer Ausführungsform ist das von dem Nukleinsäuremolekül
kodierte Protein zu mindestens etwa 50 %, vorzugsweise mindestens
etwa 60 % und stärker bevorzugt mindestens etwa 70 %, 80 % oder
40 90 % und am stärksten bevorzugt mindestens etwa 95 %, 96 %, 97 %,
98 %, 99 % oder mehr homolog zu einer Aminosäuresequenz der
Sequenz SEQ ID NO: 2, 4, 6 oder 12. Bei einer weiteren bevor-
zugten Ausführungsform ist das Protein ein Volllängen-Protein,
das im wesentlichen in Teilen homolog zu einer gesamten Amino-
45 säuresequenz der SEQ ID NO: 2, 4, 6 oder 12 (die von dem in

SEQ ID NO: 1, 3, 5 oder 11 gezeigten offenen Leserahmen herrüht) ist.

Bei anderen Ausführungsformen umfasst die isolierte Desaturase
5 eine Aminosäuresequenz, die zu mindestens etwa 50 % homolog zu einer der Aminosäuresequenzen der SEQ ID NO: 2, 4, 6 oder 12 ist und am Stoffwechsel von zum Aufbau von Fettsäuren in einem Mikroorganismus oder einer Pflanzenzelle notwendigen Verbindungen oder am Transport von Molekülen über diese Membranen teilnehmen kann,
10 wobei desaturierte C₁₈- oder C₂₀₋₂₂-Kohlenstoffketten mit Doppelbindungen an mindestens zwei Stellen gemeint ist.

Bei einer anderen bevorzugten Ausführungsform röhrt das isolierte Nukleinsäuremolekül von Phaeodactylum tricornutum UTEX646 her
15 und kodiert ein Protein (z.B. ein Desaturasefusionsprotein), das eine biologisch aktive Domäne enthält, die zu mindestens etwa 50 % oder mehr homolog zu einer Aminosäuresequenz der Sequenz SEQ ID NR 2, 4, 6 oder 12 ist und die Fähigkeit, am Stoffwechsel von zum Aufbau von Zellmembranen von Pflanzen notwendigen Verbindungen oder am Transport von Molekülen über diese Membranen teilzunehmen, beibehält oder zumindest eine der Desaturierungsaktivitäten resultierend in PUFAs wie GLA, ALA, Dihomo-gamma Linolensäure, ARA, EPA oder DHA oder deren Vorläufermoleküle besitzt, und umfasst auch heterologe Nukleinsäuresequenzen, die 25 ein heterologes Polypeptid oder regulatorische Proteine kodieren.

Alternativ kann die isolierte Desaturase eine Aminosäuresequenz umfassen, die von einer Nukleotidsequenz kodiert wird, die an eine Nukleotidsequenz der SEQ ID NO: 1, 3, 5 oder 11 hybridisiert, z.B. unter stringenten Bedingungen hybridisiert, oder zu mindestens etwa 50 %, vorzugsweise mindestens etwa 60 %, stärker bevorzugt mindestens etwa 70 %, 80 % oder 90 % und noch stärker bevorzugt mindestens etwa 95 %, 96 %, 97 %, 98 %, 99 % oder mehr homolog dazu ist. Es ist ebenfalls bevorzugt, dass die bevorzugten Desaturaseformen ebenfalls eine der hier beschriebenen Desaturaseaktivitäten besitzen.

Bei einer anderen Ausführungsform ist das isolierte Nukleinsäuremolekül mindestens 15, 25, 50, 100, 250 oder mehr Nukleotide lang
40 und hybridisiert unter stringenten Bedingungen an ein Nukleinsäuremolekül, das eine Nukleotidsequenz der SEQ ID NO: 1, 3, 5 oder 17 umfasst. Vorzugsweise entspricht das isolierte Nukleinsäuremolekül einem natürlich vorkommenden Nukleinsäuremolekül. Stärker bevorzugt kodiert das isolierte Nukleinsäuremolekül
45 natürlich vorkommende Phaeodactylum-Desaturase oder einen biologisch aktiven Teil davon.

Eine weitere Ausführungsform der Erfindung sind Expressionskassetten, die die Expression der erfindungsgemäßen Nukleinsäuren mit den Sequenzen SEQ ID NO: 1, 3, 5 oder 11 in den verschiedenen Organismen wie pflanzliche Zellen, Geweben, Teilen von Pflanzen 5 oder ganzen Pflanzen ermöglichen.

Unter der erfindungsgemäßen Expressionskassette (= Nukleinsäurekonstrukt oder -fragment) sind die in SEQ ID NO: 32 als L1 und einem Promotor, einem Strukturgen ausgewählt aus den vorteilhaftesten Sequenzen SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 11, die sich als Ergebnis des genetischen Codes und/ oder deren funktionellen oder nicht funktionellen Derivate zu verstehen, und die Polylinder-Terminator-Polylinker-Sequenzen (= L2) SEQ ID NO: 33, SEQ ID NO: 34 oder SEQ ID NO: 35 zu verstehen. 10 Diese steuern vorteilhafterweise die Genexpression in der Wirtszelle. Diese in den Konstrukten enthaltenen regulatorischen Sequenzen sollen die gezielte Expression der Gene und der Proteineexpression ermöglichen. Dies kann beispielsweise je nach Wirtsorganismus bedeuten, dass das Gen erst nach Induktion exprimiert 15 und/oder überexprimiert wird, oder dass es sofort exprimiert und/ oder überexprimiert wird. Beispielsweise handelt es sich bei diesen regulatorischen Sequenzen um Sequenzen an die Induktoren oder Repressoren binden und so die Expression der Nukleinsäure regulieren. Zusätzlich zu diesen neuen Regulationssequenzen oder an- 20 stelle dieser Sequenzen kann die natürliche Regulation dieser Sequenzen vor den eigentlichen Strukturgenen noch vorhanden sein und gegebenenfalls genetisch verändert worden sein, so dass die natürliche Regulation ausgeschaltet und die Expression der Gene erhöht wurde. Das Genkonstrukt kann aber auch einfacher aufgebaut 25 sein, das heißt es wurden keine zusätzlichen Regulationssignale vor die Nukleinsäuresequenz oder dessen Derivate inseriert und der natürliche Promotor mit seiner Regulation wurde nicht entfernt. Stattdessen wurde die natürliche Regulationssequenz so mutiert, dass keine Regulation mehr erfolgt und/oder die Genexpres- 30 sion gesteigert wird. Diese veränderten Promotoren können in Form von Teilsequenzen (= Promotor mit Teilen der erfindungsgemäßen Nukleinsäuresequenzen) auch allein vor das natürliche Gen zur Steigerung der Aktivität gebracht werden. Das Genkonstrukt kann außerdem vorteilhafterweise auch eine oder mehrere sogenannte 35 "enhancer Sequenzen" funktionell verknüpft mit dem Promotor enthalten, die eine erhöhte Expression der Nukleinsäuresequenz ermöglichen. Auch am 3'-Ende der DNA-Sequenzen können zusätzliche vorteilhafte Sequenzen inseriert werden wie weitere regulatorische Elemente oder Terminatoren. Die Δ-5-Desaturase-/Δ-6-Desatu- 40 rase und/oder Δ-12-Desaturasegene können in einer oder mehreren 45

Kopien in der Expressionskassette (= Genkonstrukt) enthalten sein.

Die regulatorischen Sequenzen bzw. Faktoren können dabei wie oben beschrieben vorzugsweise die Genexpression der eingeführten Gene positiv beeinflussen und dadurch erhöhen. So kann eine Verstärkung der regulatorischen Elemente vorteilhafterweise auf der Transkriptionsebene erfolgen, indem starke Transkriptionssignale wie Promotoren und/oder "Enhancer" verwendet werden. Daneben ist aber auch eine Verstärkung der Translation möglich, indem beispielsweise die Stabilität der mRNA verbessert wird.

Ein weiterer Aspekt der Erfindung betrifft Vektoren, z.B. rekombinante Expressionsvektoren, die mindestens eine der erfundungsgemäßen Expressionskassetten enthalten, und Wirtszellen, in die die erfundungsgemäßen Expressionskassetten oder diese Vektoren eingebracht worden sind, insbesondere Mikroorganismen, Pflanzenzellen, Pflanzengewebe, -organe oder ganze Pflanzen. Bei einer Ausführungsform kann eine solche Wirtszelle Feinchemikalien-Verbindungen, insbesondere PUFAs, speichern; zur Isolation der gewünschten Verbindung werden die Zellen geerntet. Die Verbindung (Öle, Lipide, Triacylglyceride, Fettsäuren) oder die Desaturase können dann aus dem Medium oder der Wirtszelle, welche bei Pflanzen Zellen sind, die Feinchemikalien enthalten oder speichern, am stärksten bevorzugt Zellen von Speichergeweben, wie Samenhüllen, Knollen, Epidermis- und Samenzellen, Endosperm oder Embryogewebe isoliert werden.

Noch ein weiterer Aspekt der Erfindung betrifft eine genetisch veränderte transgene Pflanze, bevorzugt ein Ölfruchtpflanze, wie vorstehend erwähnt, besonders bevorzugt eine Raps- oder Leinpflanze, in die eine erfundungsgemäße Expressionskassette, die vorteilhaft weitere Gene wie Desaturasegen enthält, eingebracht worden ist. Bei einer Ausführungsform ist das Genom von Raps oder Lein durch Einbringen einer erfundungsgemäßen Expressionskassette vorteilhaft enthaltend weitere Nukleinsäuremoleküle, die beispielsweise eine Wildtyp- oder mutierte Desaturasesequenz kodiert, als Transgen verändert worden. Bei einer bevorzugten Ausführungsform wird Raps oder Lein auch zur Produktion einer gewünschten Verbindung, wie Lipiden und Fettsäuren, wobei PUFAs besonders bevorzugt sind, verwendet.

Bei noch einer weiteren bevorzugten Ausführungsform kann das Moos *Physcomitrella patens* zur Demonstration der Funktion einer Expressionskassette mit einem Desaturasesegen unter Verwendung

21

homologer Rekombination auf der Basis der in dieser Erfindung beschriebenen Nukleinsäuren verwendet werden.

Das Desaturasepolypeptid oder ein biologisch aktiver Teil davon kann vorteilhaft unter Kontrolle der erfindungsgemäßen Expressionskassette funktionsfähig mit einem weiteren Polypeptid, das eine andere enzymatische Aktivität als die Desaturasen hat beispielsweise eine Elongase-, Acyltransferase- oder sonstige Aktivität verbunden werden, so dass ein Fusionsprotein gebildet wird.

Vorteilhaft hat dieses Fusionsprotein eine Aktivität, die sich von derjenigen der Desaturase allein unterscheidet. Bei anderen bevorzugten Ausführungsform nimmt dieses Fusionsprotein am Stoffwechsel von Verbindungen, die zur Synthese von Lipiden und Fettsäuren, Cofaktoren und Enzymen in Mikroorganismen oder Pflanzen notwendig sind, oder am Transport von Molekülen über diese Membranen teil. Besonders vorteilhaft moduliert das Einbringen dieses Fusionsproteins in einer Wirtszelle die Produktion einer gewünschten Verbindung innerhalb einer und durch die Zelle. Bei einer bevorzugten Ausführungsform enthalten diese Fusionsproteine auch Δ-4-, Δ-5- oder Δ-6, Δ-8-, Δ-15-, Δ-17 oder Δ-19-Desaturaseaktivitäten allein oder in Kombination. Insbesondere solche Genkombinationen sind bevorzugte Ausführungsformen, die aus SEQ ID NO: 7 oder 9 gewählt sind, bzw Teilen davon, Derivate oder ihren Homologen. Insbesondere solche Kombinationen sind bevorzugt, die die vollständige Proteinaktivität wie in SEQ ID NO: 1, 3, 5 oder 11 enthalten und in Multiexpressionskassetten definiert durch SEQ ID NO: 13, 14, 15, 16 und 17 eingefügt zur Transformation von Pflanzen und Expression in Pflanzen geeignet sind.

30

Eingehende Beschreibung der Erfindung

Ein erfindungsgemäßer Gegenstand sind auch die Expressionskassetten in Verbindung mit isolierten Nukleinsäuresequenz(en), die für ein Polypeptid mit Desaturaseaktivität codiert, ausgewählt aus der Gruppe:

- a) einer Nukleinsäuresequenz mit der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 11 dargestellten Sequenz,
- b) Nukleinsäuresequenzen, die aufgrund des degenerierten genetischen Codes durch Rückübersetzung der in SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6 oder SEQ ID NO: 12 dargestellten Aminosäuresequenzen erhalten werden,

c) Derivate der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 11 dargestellten Nukleinsäuresequenz, die für Polypeptide mit der in SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6 oder SEQ ID NO: 12 dargestellten Aminosäuresequenzen codieren und mindestens 50 % Homologie auf Aminosäureebene aufweisen, ohne daß die enzymatische Wirkung der Polypeptide wesentlich reduziert ist.

Weiterhin betrifft die Erfindung eine Aminosäuresequenz, die 10 durch die oben genannte(n) Nukleinsäuresequenz(en) codiert werden (für die Erfindung soll der Singular den Plural und umgekehrt umfassen). Speziell betrifft die Erfindung Aminosäuresequenzen, die durch die in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 11 dargestellte Sequenz codiert werden.

15 Die vorliegende Erfindung stellt Expressionskassetten bereit, die für die Expression von Nukleinsäuren und Proteinmoleküle mit Desaturaseaktivität geeignet sind, und von Nukleinsäuren kodierend für Proteine, die am Stoffwechsel von Lipiden und Fettsäuren, 20 PUFA-Cofaktoren und Enzymen in dem Moos *Physcomitrella patens* oder am Transport lipophiler Verbindungen über Membranen beteiligt sind. Die erfindungsgemäßen Verbindungen lassen sich zur Modulation der Produktion von Feinchemikalien aus Organismen wie Pflanzen, wie Mais, Weizen, Roggen, Hafer, Triticale, Reis, Gerste, Sojabohne, Erdnuss, Baumwolle, Linum Arten wie Öl- oder Faserlein, Brassica-Arten, wie Raps, Canola und Rübsen, Pfeffer, Sonnenblume, Borretsch, Nachtkerze und Tagetes, Solanacaen-Pflanzen, wie Kartoffel, Tabak, Aubergine und Tomate, Vicia-Arten, Erbse, Maniok, Alfalfa, Buschpflanzen (Kaffee, Kakao, Tee), Salix-Arten, Bäume (Ölpalme, Kokosnuss) und ausdauernden Gräsern und Futterfeldfrüchten, entweder direkt (z.B. wenn die Überexpression oder Optimierung eines Fettsäurebiosynthese-Proteins einen direkten Einfluss auf die Ausbeute, Produktion und/oder Effizienz der Produktion der Fettsäure aus modifizierten Organismen 35 hat) verwenden oder können eine indirekt Auswirkung haben, die dennoch zu einer Steigerung der Ausbeute, Produktion und/oder Effizienz der Produktion einer gewünschten Verbindung oder einer Abnahme unerwünschter Verbindungen führt (z.B. wenn die Modulation des Stoffwechsels von Lipiden und Fettsäuren, Cofaktoren und 40 Enzymen zu Veränderungen der Ausbeute, Produktion und/oder Effizienz der Produktion oder der Zusammensetzung der gewünschten Verbindungen innerhalb der Zellen führt, was wiederum die Produktion einer oder mehrerer Feinchemikalien beeinflussen kann). Aspekte der Erfindung sind nachstehend weiter erläutert.

45

I. Feinchemikalien und PUFAs

Der Begriff "Feinchemikalie" ist im Fachgebiet bekannt und umfasst Moleküle, die durch einen Organismus produziert worden sind und Anwendungen in verschiedenen Industrien finden, wie, aber nicht beschränkt auf, die pharmazeutische, Landwirtschafts-, Nahrungsmittel- und Kosmetik-Industrie. Diese Verbindungen umfassen Lipide, Fettsäuren, Cofaktoren und Enzyme usw. (wie z.B. beschrieben in Kuninaka, A. (1996) Nucleotides and related compounds, S. 561-612, in Biotechnology Bd. 6, Rehm et al., Hrsgb., VCH: Weinheim und darin enthaltenen Literaturstellen), Lipide, gesättigte und ungesättigte Fettsäuren (z.B. Arachidonsäure), Vitamine und Cofaktoren (wie beschrieben in Ullmann's Encyclopedia of Industrial Chemistry, Bd. A27, Vitamins, S. 443-613 (1996) VCH: Weinheim und darin enthaltenen Literaturstellen; und Ong, A.S., Niki, E., & Packer, L. (1995) Nutrition, Lipids, Health and Disease Proceedings of the UNESCO/Confederation of Scientific and Technological Associations in Malaysia and the Society for Free Radical Research - Asien, abgehalten am 1.-3. Sept. 1994 in Penang, Malaysia, AOCS Press (1995)), Enzyme und sämtliche anderen von Gutcho (1983) in Chemicals by Fermentation, Noyes Data Corporation, ISBN: 0818805086, und darin angegebenen Literaturstellen beschriebenen Chemikalien. Der Stoffwechsel und die Verwendungen bestimmter Feinchemikalien sind nachstehend weiter erläutert.

Die Kombination verschiedener Vorläufermoleküle und Biosyntheseenzyme führt zur Herstellung verschiedener Fettsäuremoleküle, was eine entscheidende Auswirkung auf die Zusammensetzung der Membran hat. Es kann angenommen werden, dass PUFAs nicht nur einfach in Triacylglycerin, sondern auch in Membranlipide eingebaut werden.

Vorläufer für die PUFA-Biosynthese sind beispielsweise Ölsäure, Linol- und Linolensäure. Diese C₁₈-Kohlenstoff-Fettsäuren müssen auf C₂₀ und C₂₂ verlängert werden, damit Fettsäuren vom Eicos- und Docos-Kettentyp erhalten werden. Mithilfe verschiedener Desaturasen, wie Enzymen, welche Δ-12-Desaturase, Δ-15-Desaturase, Δ-6-Desaturase-, Δ-5- und Δ-4-Desaturaseaktivität aufweisen, können Arachidonsäure, Eicosapentaensäure und Docosahexaensäure sowie verschiedene andere langkettige PUFAs erhalten, extrahiert und für verschiedene Zwecke bei Nahrungsmittel-, Futter-, Kosmetik- oder pharmazeutischen Anwendungen verwendet werden.

Zur Herstellung langkettiger PUFAs müssen, wie oben erwähnt, die mehrfach ungesättigten C₁₈- bzw C₂₀-Fettsäuren mehrfach desaturiert werden. Die erfindungsgemäßen Nukleinsäuresequenzen kodieren erste funktionell aktive Desaturasen aus Phyeodactylum tricornutum, einem Mikroorganismus, der PUFAs in der Triacyl-

glycerolfraktion enthält. Mit den erfindungsgemäßen Desaturasen können Doppelbindungen in die Δ-5-, Δ-6- oder Δ-12-Position eingeführt werden. Die Aktivitäten der erfindungsgemäßen Desaturasen führt vorzugsweise zu C₁₈- + C₂₀-Fettsäuren mit mindestens zwei, 5 drei, vier oder fünf Doppelbindungen im Fettsäuremolekül, vorzugsweise zu C₂₀-Fettsäuren mit vorteilhaft drei, vier oder fünf Doppelbindungen im Fettsäuremolekül. Die Desaturierung kann vor oder nach Elongation der entsprechenden Fettsäure erfolgen. Daher führen die Produkte der Desaturaseaktivitäten und der möglichen weiteren Desaturierung und Elongation zu bevorzugten PUFAAs 10 mit höherem Desaturierungsgrad, einschließlich einer weiteren Elongation von C₂₀ zu C₂₂-Fettsäuren, zu Fettsäuren wie Linolsäure, Docosadiensäure, dihomo-γ-linolensäure, Arachidonsäure, ω6-Eicosatriendihomo-γ-linolensäure, Eicosapentaensäure, ω3-Eicosatrien- 15 säure, ω3-Eicosatetraensäure, Docosapentaensäure oder Docosahexaensäure. Substrate dieser erfindungsgemäßen Enzymaktivität sind zum Beispiel Taxolsäure, 6,9-Octadecadiensäure, Ölsäure, Linolsäure, γ-Linolensäure, Pinolensäure, α-Linolensäure, Arachidonsäure, Eicosapentaensäure oder Stearidonsäure. Bevorzugte 20 Substrate sind Linolsäure, γ-Linolensäure und/oder α-Linolensäure, dihomo-γ-linolensäure bzw. Arachidonsäure, Eicosatetraensäure oder Eicosapentaensäure. Die C₁₈-oder C₂₀-Fettsäuren mit mindestens zwei Doppelbindungen in der Fettsäure können durch die erfindungsgemäße Enzymaktivität in Form der freien Fettsäure oder 25 in Form der Ester, wie Phospholipide, Glykolipide, Sphingolipide, Phosphoglyceride, Monoacylglyceride, Diacylglyceride, Triacylglyceride oder sonstige Ester, verlängert werden. Ferner müssen Fettsäuren anschließend an verschiedene Modifikationsorte transportiert und in das Triacylglycerin- 30 Speicherlipid eingebaut werden. Ein weiterer wichtiger Schritt bei der Lipidsynthese ist der Transfer von Fettsäuren auf die polaren Kopfgruppen, beispielsweise durch Glycerin-Fettsäure-Acyltransferase (siehe Frentzen, 1998, Lipid, 100(4-5):161-166). 35 Veröffentlichungen über die Pflanzen-Fettsäurebiosynthese, Desaturierung, den Lipidstoffwechsel und Membrantransport von fetthaltigen Verbindungen, die Betaoxidation, Fettsäuremodifikation und Cofaktoren, Triacylglycerin-Speicherung und -Assemblierung einschließlich der Literaturstellen darin siehe 40 in den folgenden Artikeln: Kinney, 1997, Genetic Engineering, Hrsg.: JK Setlow, 19:149-166; Ohlrogge und Browse, 1995, Plant Cell 7:957-970; Shanklin und Cahoon, 1998, Annu. Rev. Plant Physiol. Plant Mol. Biol. 49:611-641; Voelker, 1996, Genetic Engineering, Hrsg.: JK Setlow, 18:111-13; Gerhardt, 1992, Prog. 45 Lipid R. 31:397-417; Gühnemann-Schäfer & Kindl, 1995, Biochim. Biophys Acta 1256:181-186; Kunau et al., 1995, Prog. Lipid Res. 34:267-342; Stymne et al., 1993, in: Biochemistry and Molecular

Biology of Membrane and Storage Lipids of Plants, Hrsgb.: Murata und Somerville, Rockville, American Society of Plant Physiologists, 150-158, Murphy & Ross 1998, Plant Journal. 13(1):1-16.

5

Vitamine, Cofaktoren und Nutraceutical wie PUFAs, umfassen eine Gruppe von Molekülen, die höhere Tiere nicht mehr synthetisieren können und somit aufnehmen müssen oder die höhere Tiere nicht mehr ausreichend selbst herstellen können und somit zusätzlich 10 aufnehmen müssen, obwohl sie leicht von anderen Organismen, wie Bakterien, synthetisiert werden. Die Biosynthese dieser Moleküle in Organismen, die sie produzieren können, wie in Bakterien, ist im großen und ganzen charakterisiert worden (Ullmann's Encyclopedia of Industrial Chemistry, "Vitamins", Bd. A27, 15 S. 443-613, VCH: Weinheim, 1996; Michal, G. (1999) Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology, John Wiley & Sons; Ong, A.S., Niki, E., & Packer, L. (1995) "Nutrition, Lipids, Health and Disease" Proceedings of the UNESCO/Confederation of Scientific and Technological Associations 20 in Malaysia and the Society for Free Radical Research Asia, abgehalten am 1.-3. Sept. 1994 in Penang, Malaysia, AOCS Press, Champaign, IL X, 374 S.).

Die oben erwähnten Moleküle sind entweder selbst biologisch 25 aktive Moleküle oder Vorstufen biologisch aktiver Substanzen, die entweder als Elektronenüberträger oder Zwischenprodukte bei einer Vielzahl von Stoffwechselwegen dienen. Diese Verbindungen haben neben ihrem Nährwert auch einen signifikanten industriellen Wert als Farbstoffe, Antioxidantien und Katalysatoren oder andere Ver- 30 arbeitungshilfsstoffe. (Einen Überblick über Struktur, Aktivität und industrielle Anwendungen dieser Verbindungen siehe z.B. in Ullmann's Encyclopedia of Industrial Chemistry, "Vitamins", Bd. A27, S. 443-613, VCH: Weinheim, 1996). Mehrfach ungesättigte Fettsäuren haben verschiedene Funktionen und gesundheitsfördernde 35 Wirkungen, beispielsweise bei koronarer Herzerkrankung, Entzündungsmechanismen, Kinderernährung usw. Veröffentlichungen und Literaturstellen, einschließlich darin zitierter Literaturstellen, siehe in: Simopoulos, 1999, Am. J. Clin. Nutr. 70 (3. Suppl.):560-569, Takahata et al., Biosc. Biotechnol. Biochem. 40 1998, 62(11):2079-2085, Willich und Winther, 1995, Deutsche Medizinische Wochenschrift 120(7):229ff.

II. Elemente und Verfahren der Erfindung

45 Die vorliegende Erfindung beruht unter anderem auf der Entdeckung neuer Moleküle, die hier als Desaturasenukleinsäure- und -proteinmoleküle bezeichnet werden, welche eine Wirkung auf

die Produktion von Zellmembranen und Lipiden Phaeodactylum tricornutum ausüben und beispielsweise die Bewegung von Molekülen über diese Membranen beeinflussen. Bei einer Ausführungsform nehmen die Desaturasemoleküle am Stoffwechsel von zum Aufbau von 5 Zellmembranen in Organismen, wie Mikroorganismen und Pflanzen, notwendigen Verbindungen teil oder beeinflussen indirekt den Transport von Molekülen über diese Membranen. Bei einer bevorzugten Ausführungsform hat die Aktivität der erfindungsgemäßen Desaturasemoleküle zur Regulation der Produktion von Membran- 10 komponenten und des Membrantransports eine Auswirkung auf die Produktion der gewünschten Feinchemikalie durch diesen Organismus. Bei einer besonders bevorzugten Ausführungsform ist die Aktivität der erfindungsgemäßen Desaturasemoleküle moduliert, so dass die Ausbeute, Produktion und/oder Effizienz 15 der Produktion der Stoffwechselwege von Mikroorganismen oder Pflanzen, welche die erfindungsgemäßen Desaturasen regulieren, moduliert sind und die Effizienz des Transport von Verbindungen durch die Membranen verändert ist, was entweder direkt oder indirekt die Ausbeute, Produktion und/oder Effizienz der 20 Produktion einer gewünschten Feinchemikalie durch Mikroorganismen und Pflanzen moduliert.

Der Begriff "Desaturase" oder "Desaturasepolypeptid" umfasst Proteine, die an der Desaturierung von Fettsäuren teilnehmen. 25 Beispiele für Desaturasen sind in der SEQ ID NO: 1, 3, 5, 11 oder ihren Homologen, Derivaten oder Analoga offenbart. Die Begriffe Desaturase oder Desaturasenukleinsäuresequenz(en) umfassen Nukleinsäuresequenzen, die eine Desaturase kodieren und bei denen ein Teil eine kodierende Region und ebenfalls 30 entsprechende 5'- und 3'-untranslatierte Sequenzbereiche sein können. Beispiele für Desaturase-Gene sind die in SEQ ID NO: 1, 3, 5 oder 11 dargestellten. Die Begriffe Produktion oder Produktivität sind im Fachgebiet bekannt und beinhalten die Konzentration des Fermentationsproduktes (zum Beispiel der 35 gewünschten Feinchemikalie), das in einer bestimmten Zeitspanne und einem bestimmten Fermentationsvolumen gebildet wird (z.B. kg Produkt pro Stunde pro Liter). Der Begriff Effizienz der Produktion umfasst die Zeit, die zur Erzielung einer bestimmten Produktionsmenge nötig ist (z.B. wie lange die Zelle zur Auf- 40 richtung einer bestimmten Durchsatzrate einer Feinchemikalie benötigt). Der Begriff Ausbeute oder Produkt/Kohlenstoff-Ausbeute ist im Fachgebiet bekannt und umfasst die Effizienz der Umwandlung der Kohlenstoffquelle in das Produkt (d.h. die Feinchemikalie). Dies wird gewöhnlich beispielsweise ausgedrückt 45 als kg Produkt pro kg Kohlenstoffquelle. Durch Erhöhen der Ausbeute oder Produktion der Verbindung wird die Menge der gewonnenen Moleküle oder der geeigneten gewonnenen Moleküle

dieser Verbindung in einer bestimmten Kulturmenge über einen festgelegten Zeitraum erhöht. Die Begriffe Biosynthese oder Biosyntheseweg sind im Fachgebiet bekannt und umfassen die Synthese einer Verbindung, vorzugsweise einer organischen Verbindung, durch eine Zelle aus Zwischenverbindungen, beispielsweise in einem Mehrschritt- und stark regulierten Prozess. Die Begriffe Abbau oder Abbaupfad sind im Fachgebiet bekannt und umfassen die Spaltung einer Verbindung, vorzugsweise einer organischen Verbindung, durch eine Zelle in Abbauprodukte (allgemeiner gesagt, kleinere oder weniger komplexe Moleküle) beispielsweise in einem Mehrschritt- und stark regulierten Prozess. Der Begriff Stoffwechsel ist im Fachgebiet bekannt und umfasst die Gesamtheit der biochemischen Reaktionen, die in einem Organismus stattfinden. Der Stoffwechsel einer bestimmten Verbindung (z.B. der Stoffwechsel einer Fettsäure) umfasst dann die Gesamtheit der Biosynthese-, Modifikations- und Abbaupfade dieser Verbindung in der Zelle, die diese Verbindung betreffen.

Bei einer anderen Ausführungsform können die erfindungsgemäßen Nukleinsäuresequenzen, die für Desaturase-Moleküle codieren, die Produktion eines gewünschten Moleküls, wie einer Feinchemikalie, in einem Mikroorganismus oder in Pflanzen modulieren. Es gibt eine Reihe von Mechanismen, durch die die Veränderung einer erfindungsgemäßen Sequenz die Ausbeute, Produktion und/oder Effizienz der Produktion einer Feinchemikalie aus einem Mikroorganismus- oder Pflanzenstamm, die dieses veränderte Protein enthalten, direkt beeinflussen kann. Die Anzahl oder Aktivität von Desaturasen, die am Transport von Feinchemikalienmolekülen innerhalb oder aus der Zelle beteiligt sind, kann erhöht werden, so dass größere Mengen dieser Verbindungen über Membranen transportiert werden, aus denen sie leichter gewonnen und ineinander umgewandelt werden. Ferner sind Fettsäuren, Triacylglycerine und/oder Lipide selbst wünschenswerte Feinchemikalien; durch Optimierung der Aktivität oder Steigern der Anzahl einer oder mehrerer erfindungsgemäßer Desaturasen, die an der Biosynthese dieser Verbindungen beteiligt sind, oder durch Stören der Aktivität einer oder mehrerer Desaturasen, die am Abbau dieser Verbindungen beteiligt sind, kann es möglich sein, die Ausbeute, Produktion und/oder Effizienz der Produktion von Fettsäure- und Lipidmolekülen aus Organismen, wie Mikroorganismen oder Pflanzen, zu erhöhen.

Die Mutagenese der genannten Nukleinsäuresequenzen kann Desaturasen mit veränderten Aktivitäten hervorbringen, welche die Produktion einer oder mehrerer gewünschter Feinchemikalien aus Mikroorganismen oder Pflanzen indirekt beeinflussen. Beispielsweise können Desaturasen, die am Export von Abfallprodukten beteiligt

sind, eine größere Anzahl oder höhere Aktivität aufweisen, so dass die normalen Stoffwechselabfälle der Zelle (deren Menge möglicherweise aufgrund der Überproduktion der gewünschten Feinchemikalié erhöht ist) effizient exportiert werden, bevor sie die 5 Moleküle in der Zelle schädigen können (was die Lebensfähigkeit der Zelle herabsetzen würde) oder die Feinchemikalien-Biosynthese wege stören können (was die Ausbeute, Produktion oder Effizienz der Produktion einer gewünschten Feinchemikalié senken würde).

Die relativ großen intrazellulären Mengen der gewünschten Feinchemikalié selbst können ferner für die Zelle toxisch sein, so 10 dass man durch Steigern der Aktivität oder Anzahl von Transportern, die diese Verbindungen aus der Zelle exportieren können, die Lebensfähigkeit der Zelle in Kultur steigern kann, was wiederum zu einer größeren Anzahl an Zellen in der Kultur führt,

15 welche die gewünschte Feinchemikalié produzieren. Die Desaturasen können auch so manipuliert werden, dass die entsprechenden Mengen unterschiedlicher Lipid- und Fettsäuremoleküle produziert werden. Dies kann eine erhebliche Auswirkung auf die Lipidzusammensetzung der Zellmembran haben. Da jeder Lipidtyp unterschiedliche physi- 20 kalische Eigenschaften aufweist, kann eine Veränderung der Lipidzusammensetzung einer Membran die Membranfluidität signifikant verändern. Änderungen der Membranfluidität können den Transport von Molekülen über die Membran sowie die Integrität der Zelle beeinflussen, was jeweils eine erhebliche Auswirkung auf die 25 Produktion von Feinchemikalien aus Mikroorganismen und Pflanzen in Fermentationskultur im großen Maßstab hat. Pflanzenmembranen verleihen spezifische Eigenschaften, wie Toleranz gegenüber Wärme, Kälte, Salz, Trockenheit sowie Toleranz gegen Pathogene, wie Bakterien und Pilze.

30 Die genannten isolierten Nukleinsäuresequenzen sind im Genom eines *Phaeodactylum tricornutum* UTEX646-Stammes enthalten, der über die Algensammlung der University of Texas, Austin verfügbar ist.

35 Die Nukleotidsequenz der *Phaeodactylum tricornutum*-cDNA und die abgeleiteten Aminosäuresequenzen der Desaturasen sind in den SEQ ID NO: 1 bis 6 sowie 11 und 12 gezeigt. Es wurden Computeranalysen durchgeführt, die diese Nukleotidsequenzen als Sequenzen klassifizieren und/oder identifizieren, die am Stoffwechsel von 40 Zellmembrankomponenten beteiligte Proteine oder am Transport von Verbindungen über Zellmembranen beteiligte Proteine bzw. der PUFA Biosynthese codieren. EST's mit der Datenbankeingabe-NO: PT001070010R und PT001078032R durch die Erfinder stellen die erfundungsgemäßen Sequenzen in SEQ ID NO: 1 und 3 dar. Die Sequenz 45 des Fragments aus EST PT001070010R wurde ermittelt und ist wie dargestellt in SEQ ID NO: 5. Analog ist die Sequenz des Klones PT001078032R dargestellt in SEQ ID NO: 1. Den Klonen wurden Gen-

namen zugewiesen. Abkürzungen bedeuten: Pp = Physcomitrella patens, Pt = Phaeodactylum tricornutum. PT001070010R aus SEQ ID NO: 5 codiert für ein neues Gen homolog zu Δ-12-Desaturase und PT001078032R codiert für eine neuartige Δ-5-Desaturase. Pt_des6 5 kann gemäß Beispiel 5a mittels Polymerase Kettenreaktion unter Zuhilfenahme degenerierter Oligonukleotide isoliert werden. Ein so erhaltenes Fragment kann zum Sichten einer cDNA Bank aus Phaeodactylum tricornutum isoliert werden und die codierende Region einer Phaeodactylum tricornutum Δ-6-Desaturase erhalten werden. Ein so isoliertes Gen wird in Tabelle 1 als Pt_des6 bezeichnet und ist in SEQ ID NO: 3 dargestellt. Die korrespondierenden Aminosäuresequenzen werden durch Übersetzung des genetischen Codes der Sequenz ID NO: 1, 3 und 5 erhalten und sind als SEQ ID NO: 2, 4 und 6 definiert (siehe auch Tabelle 1). Auch eine weitere Nukleinsäuresequenz, die für eine Δ-12-Desaturase codiert, ist Tabelle 1 zu entnehmen. Sie trägt die Klon-Nummer PT001072031R.

Tabelle 1

20

	Genname	Klonname	Nukleinsäure SEQ ID NO:	Polypeptid SEQ ID NO:
D5 Desaturase	Pt_des5	PT001078032R	1	2
D6 Desaturase	Pt_des6	Pt_des6	3	4
25 D12 Desaturase	Pt_des12	PT001070010R	5	6
D6 Desaturase	Pp_des6	Pp_des6	7	8
D6 Elongase	Pp_PSE1	PP001019019F	9	10
Δ12 Desaturase	Pt_des12.2	PT001072013R	11	12

30 Die vorliegende Erfindung betrifft auch Proteine mit einer Aminosäuresequenz, die im wesentlichen homolog zu einer Aminosäuresequenz der SEQ ID NO: 2, 4, 6 oder 12 ist. Wie hier verwendet, ist ein Protein mit einer Aminosäuresequenz, die im wesentlichen homolog zu einer ausgewählten Aminosäuresequenz ist, zu mindestens etwa 50 % homolog zu der ausgewählten Aminosäuresequenz, z.B. der gesamten ausgewählten Aminosäuresequenz. Ein Protein mit einer Aminosäuresequenz, die im wesentlichen homolog zu einer ausgewählten Aminosäuresequenz ist, kann auch zu mindestens etwa 50 bis 60 %, vorzugsweise mindestens etwa 60 bis 70 % und stärker 35 bevorzugt mindestens etwa 70 bis 80 %, 80 bis 90 % oder 90 bis 95 % und am stärksten bevorzugt mindestens etwa 96 %, 97 %, 98 %, 99 % oder mehr homolog zu einer ausgewählten Aminosäuresequenz sein.

45 Die Desaturasen oder der biologisch aktive Teile oder Fragmente davon können am Stoffwechsel von Lipiden zum Aufbau von Zellmembranen oder Speicherlipiden in Organismen teilnehmen und in Kom-

30

bination mit weiteren Genen, insbesondere solchen mit Elongaseaktivität zur Elongation von C₁₈-bzw C₂₀₋₂₂-PUFAs benötigten Aktivitäten beitragen, so dass C₁₈, C₂₀-, C₂₂- oder C₂₄-PUFAs sowie verwandte PUFAs erhalten werden.

5

Dabei können die Desaturasen in Kombination mit Elongasen und anderen Desaturasen in erfindungsgemäßen Expressionskassetten kloniert werden und zur Transformation von Pflanzen mithilfe von Agrobakterium eingesetzt werden.

10

Verschiedene Aspekte der Erfindung sind eingehender in den folgenden Unterabschnitten beschrieben.

A. Isolierte Nukleinsäuremoleküle

15

Eine Ausführungsform der Erfindung sind isolierte Nukleinsäuren, die von PUFA produzierenden Mikroorganismen stammen und für Polypeptide kodieren, die C₁₈-oder C₂₀₋₂₂-Fettsäuren mit mindestens einer, zwei, drei oder vier Doppelbindungen in der Fettsäure desaturieren.

20

Eine weitere erfindungsgemäße Ausführungsform sind isolierte Nukleinsäuren, umfassend Nukleotidsequenzen, die für Polypeptide kodieren, die C₁₈-bzw C₂₀-Fettsäuren mit mindestens ein, zwei, drei oder vier Doppelbindungen in der Fettsäure desaturieren und sind aus der Gruppe, bestehend aus

- a) einer Nukleinsäuresequenz mit der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 11 dargestellten Sequenz,
- b) Nukleinsäuresequenzen, die aufgrund des degenerierten genetischen Codes durch Rückübersetzung der in SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6 oder SEQ ID NO: 12 dargestellten Aminosäuresequenzen erhalten werden,
- c) Derivate der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 11 dargestellten Nukleinsäuresequenz, die für Polypeptide mit der in SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6 oder SEQ ID NO: 12 dargestellten Aminosäuresequenzen codieren und mindestens 50 % Homologie auf Aminosäureebene aufweisen, ohne dass die enzymatische Wirkung der Polypeptide wesentlich reduziert ist.

45

Die oben genannte erfindungsgemäße Nukleinsäure stammt von Organismen, wie Ciliaten, Pilzen, Algen oder Dinoflagellaten, die PUFA-s synthetisieren können, vorzugsweise von Phaeodactylum tricornutum oder nah verwandten Organismen.

5

Ein Aspekt der Erfindung betrifft isolierte Expressionskassetten sowie Nukleinsäuremoleküle, die Desaturase-Polypeptide oder biologisch aktive Teile davon kodieren, sowie Nukleinsäurefragmente von diesen, die zur Verwendung als Hybridisierungssonden oder

10 Primer zur Identifizierung oder Amplifizierung einer Desaturase-kodierenden Nukleinsäure (z.B. Desaturase-DNA) ausreichen. Der Begriff "Nukleinsäuremolekül", wie hier verwendet, soll DNA-Moleküle (z.B. cDNA oder genomische DNA) und RNA-Moleküle (z.B. mRNA) sowie DNA- oder RNA-Analoga, die mittels Nukleotidanalogia erzeugt

15 werden, umfassen. Dieser Begriff umfasst zudem die am 3'- und am 5'-Ende des kodierenden Genbereichs gelegene untranslatierte Sequenz: mindestens 500, bevorzugt 200, besonders bevorzugt 100 Nukleotide der Sequenz stromaufwärts des 5'-Endes des kodierenden Bereichs und mindestens 100, bevorzugt 50, besonders bevorzugt 20

20 Nukleotide der Sequenz stromabwärts des 3'-Endes des kodierenden Genbereichs. Das Nukleinsäuremolekül kann einzelsträngig oder doppelsträngig sein, ist aber vorzugsweise doppelsträngige DNA. Ein "isoliertes" Nukleinsäuremolekül wird von anderen Nukleinsäuremolekülen abgetrennt, die in der natürlichen Quelle der Nukleinsäure vorliegen. Eine "isolierte" Nukleinsäure hat vorzugs-

25 weise keine Sequenzen, welche die Nukleinsäure in der genomischen DNA des Organismus, aus dem die Nukleinsäure stammt, natürlicherweise flankieren (z.B. Sequenzen, die sich an den 5'- und 3'-Enden der Nukleinsäure befinden). Bei verschiedenen Ausführungsformen kann das isolierte Desaturase-Nukleinsäuremolekül zum Beispiel weniger als etwa 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0,5 kb oder 0,1 kb an Nukleotidsequenzen enthalten, die natürlicherweise das Nukleinsäuremolekül in der genomischen DNA der Zelle, aus der die Nukleinsäure stammt (z.B. eine Physcomitrella patens-Zelle)

30 35 flankieren. Ein "isoliertes" Nukleinsäuremolekül, wie ein cDNA-Molekül, kann überdies im wesentlichen frei von anderem zellulären Material oder Kulturmedium sein, wenn es durch rekombinante Techniken hergestellt wird, oder frei von chemischen Vorstufen oder anderen Chemikalien sein, wenn es chemisch synthetisiert

40 wird.

Ein erfindungsgemäße Expressionskassette mit der Struktur SEQ ID NO: 32 -Promotor- SEQ ID NO: 33, SEQ ID NO: 34 oder SEQ ID NO: 35 Nukleinsäuremolekül, z.B. ein Nukleinsäuremolekül mit einer Nukleotidsequenz der SEQ ID NO:1 oder eines Teils davon, kann unter Verwendung molekularbiologischer Standardtechniken und der hier bereitgestellten Sequenzinformation isoliert werden. Auch kann

mithilfe von Vergleichsalgorithmen beispielsweise eine homologe Sequenz oder homologe, konservierte Sequenzbereiche auf DNA oder Aminosäureebene identifiziert werden. Beispielsweise kann aus einer *Phaeodactylum tricornutum* cDNA aus einer *Phaeodactylum tricornutum*-Bank isoliert werden, indem die vollständige SEQ ID NO:1, 3, 5 oder 11 oder ein Teil davon als Hybridisierungssonde sowie Standard-Hybridisierungstechniken (wie z.B. beschrieben in Sambrook et al., Molecular Cloning: A Laboratory Manual.

2. Aufl., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989) verwendet werden.

Überdies lässt sich ein Nukleinsäuremolekül, umfassend eine vollständige Sequenz der SEQ ID NO: 1, 3, 5 oder 11 oder einen Teil davon, durch Polymerasekettenreaktion isolieren, wobei Oligonukleotidprimer, die auf der Basis dieser Sequenz oder von Teilen davon, insbesondere Regionen um Motive aus Beispiel 5a erstellt werden oder Modifikationen ebensolcher in einzelnen definierten Aminosäuren, verwendet werden (z.B. kann ein Nukleinsäuremolekül, umfassend die vollständigen Sequenz der SEQ ID NO:1, 3, 5 oder 11 oder einen Teil davon, durch Polymerasekettenreaktion unter Verwendung von Oligonukleotidprimern isoliert werden, die auf der Basis dieser gleichen Sequenz der SEQ ID NO: 1, 3, 5 oder 11 erstellt worden sind). Zum Beispiel lässt sich mRNA aus Zellen isolieren (z.B. durch das Guanidiniumthiocyanat-Extraktionsverfahren von Chirgwin et al. (1979) Biochemistry 18:5294-5299) und cDNA mittels Reverser Transkriptase (z.B. Moloney-MLV-Reverse-Transkriptase, erhältlich von Gibco/BRL, Bethesda, MD, oder AMV-Reverse-Transkriptase, erhältlich von Seikagaku America, Inc., St.Petersburg, FL) herstellen. Synthetische Oligonukleotidprimer zur Amplifizierung mittels Polymerasekettenreaktion lassen sich auf der Basis einer der in SEQ ID NO: 1, 3, 5 oder 11 sowie der in Figur 5a gezeigten Sequenzen oder mithilfe der in SEQ ID NO: 2, 4, 6 oder 12 dargestellten Aminosäuresequenzen erstellen. Eine erfindungsgemäße Nukleinsäure kann unter Verwendung von cDNA oder alternativ von genomischer DNA als Matrize und geeigneten Oligonukleotidprimern gemäß Standard-PCR-Amplifikationstechniken amplifiziert werden. Die so amplifizierte Nukleinsäure kann in einen geeigneten Vektor kloniert werden und mittels DNA-Sequenzanalyse charakterisiert werden. Oligonukleotide, die einer Desaturase-Nukleotidsequenz entsprechen, können durch Standard-Syntheseverfahren, beispielsweise mit einem automatischen DNA-Synthesegerät, hergestellt werden.

Die in SEQ ID NO: 1,3, 5 oder 11 gezeigte cDNA umfasst Sequenzen, die Desaturasen kodieren, (d.h. den "kodierenden Bereich") sowie 5'-untranslatierte Sequenzen und 3'-untranslatierte Sequenzen. Alternativ kann das Nukleinsäuremolekül nur den kodierenden Bereich einer der Sequenzen in SEQ ID NO: 1, 3, 5 oder 11 um-

fassen oder kann ganze genomische Fragmente, die aus genomischer DNA isoliert sind, enthalten.

Bei einer weiteren bevorzugten Ausführungsform umfasst ein erfindungsgemäßes isoliertes Nukleinsäuremolekül ein Nukleinsäuremolekül, das ein Komplement einer der in SEQ ID NO: 1, 3, 5 oder 11 gezeigten Nukleotidsequenzen oder eines Teils davon ist. Ein Nukleinsäuremolekül, das zu einer der in SEQ ID NO: 1, 3, 5 oder 11 gezeigten Nukleotidsequenzen komplementär ist, ist dann ausreichend komplementär, wenn es mit einer der in SEQ ID NO: 1, 3, 5 oder 11 angegebenen Sequenzen hybridisieren kann, wodurch ein stabiler Duplex entsteht.

Homologe der neuen Desaturase-Nukleinsäuresequenzen mit der Sequenz SEQ ID NO: 1, 3, 5 oder 11 bedeutet beispielsweise allelische Varianten mit mindestens etwa 50 bis 60 %, vorzugsweise mindestens etwa 60 bis 70 %, stärker bevorzugt mindestens etwa 70 bis 80 %, 80 bis 90 % oder 90 bis 95 % und noch stärker bevorzugt mindestens etwa 95 %, 96 %, 97 %, 98 %, 99 % oder mehr Homologie zu einer in SEQ ID NO: 1, 3, 5 oder 11 gezeigten Nukleotidsequenzen oder ihren Homologen, Derivaten oder Analoga oder Teilen davon. Bei einer weiteren bevorzugten Ausführungsform umfasst ein isoliertes erfindungsgemäßes Nukleinsäuremolekül eine Nukleotidsequenz, die an eine der in SEQ ID NO: 1, 3, 5 oder 11 gezeigten Nukleotidsequenzen oder einen Teil davon hybridisiert, z.B. unter stringenten Bedingungen hybridisiert. Allelische Varianten umfassen insbesondere funktionelle Varianten, die sich durch Deletion, Insertion oder Substitution von Nukleotiden aus/in der in SEQ ID NO: 1, 3, 5 oder 11 dargestellten Sequenz erhalten lassen, wobei aber die Absicht ist, dass die Enzymaktivität der davon herrührenden synthetisierten Proteine für die Insertion eines oder mehrerer Gene vorteilhafterweise beibehalten wird. Proteine, die noch die enzymatische Aktivität der Desaturase besitzen, das heißt deren Aktivität im wesentlichen nicht reduziert ist, bedeutet Proteine mit mindestens 10 %, vorzugsweise 20 %, besonders bevorzugt 30 %, ganz besonders bevorzugt 40 % der ursprünglichen Enzymaktivität, verglichen mit dem durch SEQ ID NO: 2, 4, 6 oder 12 kodierten Protein.

Homologen der SEQ ID NO: 1, 3, 5 oder 11 bedeuten beispielsweise auch bakterielle, Pilz- und Pflanzenhomologen, verkürzte Sequenzen, einzelsträngige DNA oder RNA der kodierenden und nicht-kodierenden DNA-Sequenz.

Homologen der SEQ ID NO: 1, 3, 5 oder 11 bedeutet auch Derivate, wie beispielsweise Promotorvarianten. Die Promotoren stromaufwärts der angegebenen Nukleotidsequenzen können durch

einen oder mehrere Nukleotidaustausche, durch Insertion(en) und/oder Deletion(en) modifiziert werden, ohne dass jedoch die Funktionalität oder Aktivität der Promotoren gestört wird. Es ist weiterhin möglich, dass die Aktivität der Promotoren durch 5 Modifikation ihrer Sequenz erhöht ist oder dass sie vollständig durch aktiveren Promotoren, sogar aus heterologen Organismen, ersetzt werden.

Überdies kann das erfindungsgemäße Nukleinsäuremolekül nur 10 einen Teil des kodierenden Bereichs einer der Sequenzen in SEQ ID NO: 1, 3, 5 oder 11 umfassen, zum Beispiel ein Fragment, das als Sonde oder Primer verwendet werden kann, oder ein Fragment, welches einen biologisch aktiven Abschnitt einer Desaturase kodiert. Die aus der Klonierung des Desaturase-Gens 15 von Phaeodactylum tricornutum ermittelten Nukleotidsequenzen ermöglichen die Erzeugung von Sonden und Primern, die zur Identifizierung und/oder Klonierung von Desaturase-Homologen in anderen Zelltypen und Organismen sowie Desaturase-Homologen aus anderen Mikroalgen oder verwandten Arten gestaltet sind. Die Sonde/der 20 Primer umfasst gewöhnlich im wesentlichen gereinigtes Oligonukleotid. Das Oligonukleotid umfasst gewöhnlich einen Nukleotidsequenzbereich, der unter stringenten Bedingungen an mindestens etwa 12, vorzugsweise etwa 16, stärker bevorzugt etwa 25, 40, 50 oder 75 aufeinanderfolgende Nukleotide eines Sense-Stranges einer 25 der in SEQ ID NO: 1, 3, 5 oder 11 angegebenen Sequenzen, eines Antisense-Stranges einer der in SEQ ID NO: 1, 3, 5 oder 11 angegebenen Sequenzen oder seiner Homologen, Derivate oder Analoga oder natürlich vorkommender Mutanten davon hybridisiert. Primer auf der Basis einer Nukleotidsequenz der SEQ ID NO: 1, 3, 5 30 oder 11 können in PCR-Reaktionen zur Klonierung von Desaturase-Homologen verwendet werden. Sonden auf der Basis der Desaturase-Nukleotidsequenzen können zum Nachweis von Transkripten oder genomischen Sequenzen, die das gleiche oder homologe Proteine kodieren, verwendet werden. Bei bevorzugten Ausführungsformen 35 umfasst die Sonde zudem eine daran gebundene Markierungsgruppe, z.B. ein Radioisotop, eine fluoreszierende Verbindung, ein Enzym oder einen Enzym-Cofaktor. Diese Sonden können als Teil eines Test-Kits für genomische Marker zur Identifizierung von Zellen, die eine Desaturase misexprimieren, beispielsweise durch Messen 40 einer Menge einer Desaturase-kodierenden Nukleinsäure in einer Zellenprobe, z.B. Messen der Desaturase-mRNA-Spiegel, oder zur Bestimmung, ob ein genomisches Desaturase-Gen mutiert oder deletiert ist, verwendet werden.

45 Bei einer Ausführungsform kodiert das erfindungsgemäße Nukleinsäuremolekül ein Protein oder einen Teil davon, das/der eine Aminosäuresequenz umfasst, die ausreichend homolog zu einer

Aminosäuresequenz der SEQ ID NO: 2, 4, 6 oder 12 ist, dass das Protein oder der Teil davon die Fähigkeit, am Stoffwechsel von zum Aufbau von Zellmembranen in Mikroorganismen oder Pflanzen notwendigen Verbindungen oder am Transport von Molekülen über diese Membranen teilzunehmen, beibehält. Wie hier verwendet, betrifft der Begriff "ausreichend homolog" Proteine oder Teile davon, deren Aminosäuresequenzen eine minimale Anzahl identischer oder äquivalenter Aminosäurereste (z.B. einen Aminosäurerest mit einer ähnlichen Seitenkette, wie ein Aminosäurerest in einer der Sequenzen der SEQ ID NO:2) zu einer Aminosäuresequenz der SEQ ID NO: 2 aufweisen, so dass das Protein oder der Teil davon am Stoffwechsel von zum Aufbau von Zellmembranen in Mikroorganismen oder Pflanzen notwendigen Verbindungen oder am Transport von Molekülen über diese Membranen teilnehmen kann. Proteinbestandteile dieser Stoffwechselwege für Membrankomponenten oder Membrantransportsysteme können, wie hier beschrieben, eine Rolle bei der Produktion und Sekretion einer oder mehrerer Feinchemikalien spielen. Beispiele für diese Aktivitäten sind hier ebenfalls beschrieben. Somit trägt die "Funktion einer Desaturase" entweder direkt oder indirekt zur Ausbeute, Produktion und/oder Effizienz der Produktion einer oder mehrerer Feinchemikalien bei. Beispiele für Desaturase-Substratspezifitäten der katalytischen Aktivität sind in Tabelle 5 und 6 angegeben.

25

Bei einer weiteren Ausführungsform kodieren Derivate des erfindungsgemäßen Nukleinsäuremoleküls Proteine mit mindestens etwa 50 bis 60 %, vorzugsweise mindestens etwa 60 bis 70 % und stärker bevorzugt mindestens etwa 70 bis 80 %, 80 bis 90 %, 90 bis 95 % und am stärksten bevorzugt mindestens etwa 96 %, 97 %, 98 %, 99 % oder mehr Homologie zu einer vollständigen Aminosäuresequenz der SEQ ID NO:2. Die Homologie der Aminosäuresequenz kann über den gesamten Sequenzbereich mit dem Programm PileUp (J. Mol. Evolution., 25, 351-360, 1987, Higgins et al., CABIOS, 5, 1989:151-153) oder BESTFIT oder GAP bestimmt (Henikoff, S. and Henikoff, J. G. (1992). Amino acid substitution matrices from protein blocks. Proc. Natl. Acad. Sci. USA 89: 10915-10919.)

40 Teile von Proteinen, die von den erfindungsgemäßen Desaturase-Nukleinsäuremolekülen kodiert werden, sind vorzugsweise biologisch aktive Teile einer der Desaturasen. Wie hier verwendet, soll der Begriff "biologisch aktiver Teil einer Desaturase", einen Abschnitt, z.B. eine Domäne/ein Motiv, einer Desaturase umfassen, der am Stoffwechsel von zum Aufbau von Zellmembranen in Mikroorganismen oder Pflanzen notwendigen Verbindungen oder am Transport von Molekülen über diese Membranen teilnehmen kann

oder eine in Tabelle 5 und 6 angegebene Aktivität aufweist. Zur Bestimmung, ob eine Desaturase oder ein biologisch aktiver Teil davon am Stoffwechsel von zum Aufbau von Zellmembranen in Mikroorganismen oder Pflanzen notwendigen Verbindungen oder am Transport von Molekülen über diese Membranen teilnehmen kann, kann ein Test der enzymatischen Aktivität durchgeführt werden. Diese Testverfahren, wie eingehend in Beispiel 8 des Beispielteils beschrieben, sind dem Fachmann geläufig.

10 Zusätzliche Nukleinsäurefragmente, die biologisch aktive Abschnitte einer Desaturase kodieren, lassen sich durch Isolierung eines Teils einer der Sequenzen in SEQ ID NO: 1, 3, 5 oder 11, Exprimieren des kodierten Abschnitt der Desaturase oder des Peptids (z.B. durch rekombinante Expression in vitro) 15 und Bestimmen der Aktivität des kodierten Teils der Desaturase oder des Peptids herstellen.

Die Erfindung umfasst zudem Nukleinsäuremoleküle, die sich von einer der in SEQ ID NO: 1, 3, 5 oder 11 gezeigten Nukleotidsequenzen (und Teilen davon) aufgrund des degenerierten genetischen Codes unterscheiden und somit die gleiche Desaturase kodieren wie diejenige, die von den in SEQ ID NO: 1, 3, 5 oder 11 gezeigten Nukleotidsequenzen kodiert wird. Bei einer anderen Ausführungsform hat ein erfindungsgemäßes isoliertes Nukleinsäuremolekül eine Nukleotidsequenz, die für ein Protein mit einer in SEQ ID NO: 2, 4, 6 oder 12 gezeigten Aminosäuresequenz kodiert. Bei einer weiteren Ausführungsform kodiert das erfindungsgemäß Nukleinsäuremolekül ein Vollängen-Desaturase-Protein, das zu einer Aminosäuresequenz der SEQ ID NO: 2, 4, 6 oder 12 (die 30 von einem in SEQ ID NO: 1, 3, 5 oder 11 gezeigten offenen Leseraster kodiert wird) im wesentlichen homolog ist und durch gängige Methoden identifizierbar und isolierbar ist.

Zusätzlich zu den in SEQ ID NO: 1, 3, 5 oder 11 gezeigten 35 Desaturase-Nukleotidsequenzen erkennt der Fachmann, dass DNA-Sequenzpolymorphismen, die zu Änderungen in den Aminosäuresequenzen der Desaturasen führen, innerhalb einer Population (z.B. der Phaeodactylum tricornutum-Population) existieren können. Diese genetischen Polymorphismen im Desaturase-Gen können 40 zwischen Individuen innerhalb einer Population aufgrund von natürlicher Variation existieren. Wie hier verwendet, bedeuten die Begriffe "Gen" und "rekombinantes Gen" Nukleinsäuremoleküle mit einem offenen Leserahmen, der eine Desaturase, vorzugsweise eine Phaeodactylum tricornutum -Desaturase, kodiert. Diese natürlichen Varianten bewirken üblicherweise eine Varianz von 1 bis 45 5 % in der Nukleotidsequenz des Desaturase-Gens. Sämtliche und alle dieser Nukleotidvariationen und daraus resultierende

Aminosäurepolymorphismen in der Desaturase, die das Ergebnis natürlicher Variation sind und die funktionelle Aktivität von Desaturasen nicht verändern, sollen im Umfang der Erfindung enthalten sein.

5

Nukleinsäuremoleküle, die den natürlichen Varianten entsprechen, und nicht-*Phaeodactylum tricornutum*-Homologen, -Derivate und -Analoga der *Phaeodactylum tricornutum*-cDNA können auf der Grundlage ihrer Homologie zu der hier offenbarten *Phaeodactylum tricornutum*-Desaturase-Nukleinsäure unter Verwendung der *Phaeodactylum tricornutum*-cDNA oder eines Teils davon als Hybridisierungssonde gemäß Standard-Hybridisierungstechniken unter stringenten Hybridisierungsbedingungen isoliert werden. Bei einer anderen Ausführungsform ist ein erfindungsgemäßes isoliertes Nukleinsäuremolekül mindestens 15 Nukleotide lang und hybridisiert unter stringenten Bedingungen mit dem Nukleinsäuremolekül, das eine Nukleotidsequenz der SEQ ID NO:1, 3, 5 oder 11 umfasst. Bei anderen Ausführungsformen ist die Nukleinsäure mindestens 25, 50, 100, 250 oder mehr Nukleotide lang. Der Begriff "hybridisiert unter stringenten Bedingungen", wie hier verwendet, soll Hybridisierungs- und Waschbedingungen beschreiben, unter denen Nukleotidsequenzen, die mindestens 60 % homolog zueinander sind, gewöhnlich aneinander hybridisiert bleiben. Die Bedingungen sind vorzugsweise derart, dass Sequenzen, die mindestens etwa 65 %, stärker bevorzugt mindestens etwa 70 % und noch stärker bevorzugt mindestens etwa 75 % oder stärker zueinander homolog sind, gewöhnlich aneinander hybridisiert bleiben. Diese stringenten Bedingungen sind dem Fachmann bekannt und lassen sich in Current Protocols in Molecular Biology, John Wiley & Sons, N. Y. (1989), 6.3.1-6.3.6., finden. Ein bevorzugtes, nicht einschränkendes Beispiel für stringenten Hybridisierungsbedingungen sind Hybridisierungen in 6 x Natriumchlorid/Natriumcitrat (sodium chloride/ sodiumcitrate = SSC) bei etwa 45°C, gefolgt von einem oder mehreren Waschschriften in 0,2 x SSC, 0,1 % SDS bei 50 bis 65°C. Dem Fachmann ist bekannt, dass diese Hybridisierungsbedingungen sich je nach dem Typ der Nukleinsäure und, wenn beispielsweise organische Lösungsmittel vorliegen, hinsichtlich der Temperatur und der Konzentration des Puffers unterscheiden. Die Temperatur unterscheidet sich beispielsweise unter "Standard-Hybridisierungsbedingungen" je nach dem Typ der Nukleinsäure zwischen 42°C und 58°C in wässrigem Puffer mit einer Konzentration von 0,1 bis 5 x SSC (pH 7,2). Falls organisches Lösungsmittel im oben genannten Puffer vorliegt, zum Beispiel 50 % Formamid, ist die Temperatur unter Standardbedingungen etwa 42°C. Vorzugsweise sind die Hybridisierungsbedingungen für DNA:DNA-Hybride zum Beispiel 0,1 x SSC und 20°C bis 45°C, vorzugsweise zwischen 30°C und 45°C. Vorzugsweise sind die Hybridisierungsbedingungen für DNA:RNA-

Hybride zum Beispiel 0,1 x SSC und 30°C bis 55°C, vorzugsweise zwischen 45°C und 55°C. Die vorstehend genannten Hybridisierungs-temperaturen sind beispielsweise für eine Nukleinsäure mit etwa 100 bp (= Basenpaare) Länge und einem G + C-Gehalt von 50 %

5 in Abwesenheit von Formamid bestimmt. Der Fachmann weiß, wie die erforderlichen Hybridisierungsbedingungen anhand von Lehrbüchern, wie dem vorstehend erwähnten oder aus den folgenden Lehrbüchern Sambrook et al., "Molecular Cloning", Cold Spring Harbor Laboratory, 1989; Hames und Higgins (Hrsgb.) 1985,

10 "Nucleic Acids Hybridization: A Practical Approach", IRL Press at Oxford University Press, Oxford; Brown (Hrsgb.) 1991, "Essential Molecular Biology: A Practical Approach", IRL Press at Oxford University Press, Oxford, bestimmt werden können.

15 Vorzugsweise entspricht ein erfindungsgemäßes isoliertes Nuklein-säuremolekül, das unter stringenten Bedingungen an eine Sequenz der SEQ ID NO:1, 3, 5 oder 11 hybridisiert, einem natürlich vor-kommenden Nukleinsäuremolekül. Wie hier verwendet, betrifft ein "natürlich vorkommendes" Nukleinsäuremolekül ein RNA- oder DNA-

20 Molekül mit einer Nukleotidsequenz, die in der Natur vorkommt (z.B. ein natürliches Protein kodiert). Bei einer Ausführungsform kodiert die Nukleinsäure eine natürliche vorkommende Phyaeo-dactylum tricornutum-Desaturase.

25 Zusätzlich zu natürlich vorkommenden Varianten der Desaturase-sequenz, die in der Population existieren können, erkennt der Fachmann ferner, dass auch Änderungen durch Mutation in einer Nukleotidsequenz der SEQ ID NO: 1, 3, 5 oder 11 eingebracht werden können, was zu Änderungen der Aminosäuresequenz der

30 kodierten Desaturase führt, ohne dass die Funktionsfähigkeit des Desaturaseproteins beeinträchtigt wird. Beispielsweise lassen sich Nukleotidsusbtitutionen, die an "nicht-essentiellen" Aminosäureresten zu Aminosäuresubstitutionen führen, in einer Sequenz der SEQ ID NO: 2, 4, 6 oder 12 herstellen. Ein "nicht-

35 essentieller" Aminosäurerest ist ein Rest, der sich in einer Wildtyp-Desaturasesequenz einer der Desaturasen (SEQ ID NO: 2, 4, 6 oder 12) verändern lässt, ohne dass die Aktivität der Desaturase verändert das heißt wesentlich reduziert wird, wohingegen ein "essentieller" Aminosäurerest für die Desaturaseaktivität

40 erforderlich ist. Andere Aminosäurereste (z.B. diejenigen, die in der Domäne mit Desaturaseaktivität nicht konserviert oder lediglich semikonserviert sind) können jedoch für die Aktivität nicht essentiell sein und lassen sich somit verändern, ohne dass die Desaturaseaktivität verändert wird.

Folglich betrifft ein weiterer Aspekt der Erfindung Nuklein-säuremoleküle, die Desaturasen kodieren, die veränderte Amino-säurereste enthalten, die für die Desaturaseaktivität nicht essentiell sind. Diese Desaturasen unterscheiden sich in der 5 Aminosäuresequenz von einer Sequenz in SEQ ID NO: 2, 4, 6 oder 12 und behalten dennoch zumindest eine der hier beschriebenen Desaturaseaktivitäten. Das isolierte Nukleinsäuremolekül umfasst bei einer Ausführungsform eine Nukleotidsequenz, die ein Protein kodiert, wobei das Protein eine Aminosäuresequenz mit mindestens 10 etwa 50 % Homologie zu einer Aminosäuresequenz der SEQ ID NO: 2, 4, 6 oder 12 umfasst und am Stoffwechsel von zum Aufbau von Zell-membranen in Phaeodactylum tricornutum notwendigen Verbindungen oder am Transport von Molekülen über diese Membranen teilnehmen kann. Das von dem Nukleinsäuremolekül kodierte Protein ist 15 vorzugsweise mindestens etwa 50 bis 60 % homolog zu einer der Sequenzen in SEQ ID NO: 2, 4, 6 oder 12 stärker bevorzugt mindestens etwa 60 bis 70 % homolog zu einer der Sequenzen in SEQ ID NO: 2, 4, 6 oder 12 noch stärker bevorzugt mindestens etwa 70 bis 80 %, 80 bis 90 %, 90 bis 95 % homolog zu einer 20 der Sequenzen in SEQ ID NO: 2, 4, 6 oder 12 und am stärksten bevorzugt mindestens etwa 96 %, 97 %, 98 % oder 99 % homolog zu einer der Sequenzen in SEQ ID NO: 2, 4, 6 oder 12.

Zur Bestimmung der prozentualen Homologie von zwei Aminosäure- 25 sequenzen (z.B. einer der Sequenzen der SEQ ID NO: 2, 4, 6 oder 12 und einer mutierten Form davon) oder von zwei Nukleinsäuren werden die Sequenzen zum Zweck des optimalen Vergleichs unter-einander geschrieben (z.B. können Lücken in die Sequenz eines Proteins oder einer Nukleinsäure eingefügt werden, um ein opti- 30 males Alignment mit dem anderen Protein oder der anderen Nuklein-säure zu erzeugen). Die Aminosäurereste oder Nukleotide an den entsprechenden Aminosäurepositionen oder Nukleotidpositionen werden dann verglichen. Wenn eine Position in einer Sequenz (z.B. einer der Sequenzen der SEQ ID NO: 2, 4, 6 oder 12) durch den 35 gleichen Aminosäurerest oder das gleiche Nukleotid wie die ent-sprechende Stelle in der anderen Sequenz (z.B. einer mutierten Form der aus SEQ ID NO: 2, 4, 6 oder 12 ausgewählten Sequenz) belegt wird, dann sind die Moleküle an dieser Position homolog (d.h. Aminosäure- oder Nukleinsäure- "Homologie", wie hier ver- 40 wendet, entspricht Aminosäure- oder Nukleinsäure- "Identität"). Die prozentuale Homologie zwischen den beiden Sequenzen ist eine Funktion der Anzahl an identischen Positionen, die den Sequenzen gemeinsam sind (d.h. % Homologie = Anzahl der identischen Positionen/Gesamtanzahl der Positionen x 100). 45 Die Begriffe Homologie und Identität sind damit als Synonym anzusehen.

40

Ein isoliertes Nukleinsäuremolekül, das eine Desaturase kodiert, die zu einer Proteinsequenz der SEQ ID NO: 2, 4, 6 oder 12 homolog ist, kann durch Einbringen einer oder mehrerer Nukleotidsubstitutionen, -additionen oder -deletionen in eine Nukleotidsequenz der SEQ ID NO: 1, 3, 5 oder 11 erzeugt werden, so dass eine oder mehrere Aminosäuresubstitutionen, -additionen oder -deletionen in das kodierte Protein eingebracht werden.

Mutationen können in eine der Sequenzen der SEQ ID NO: 1, 3, 5 oder 11 durch Standardtechniken, wie stellenspezifische Mutationen und PCR-vermittelte Mutagenese, eingebracht werden.

Vorzugsweise werden konservative Aminosäuresubstitutionen an einer oder mehreren der vorhergesagten nicht-essentiellen Aminosäureresten hergestellt. Bei einer "konservativen Aminosäuresubstitution" wird der Aminosäurerest gegen einen Aminosäurerest mit einer ähnlichen Seitenkette ausgetauscht. Im Fachgebiet sind Familien von Aminosäureresten mit ähnlichen Seitenketten definiert worden. Diese Familien umfassen Aminosäuren mit basischen Seitenketten (z.B. Lysin, Arginin, Histidin), sauren Seitenketten (z.B. Asparaginsäure, Glutaminsäure), ungeladenen polaren Seitenketten (z.B. Glycin, Asparagin, Glutamin, Serin, Threonin, Tyrosin, Cystein), unpolaren Seitenketten, (z.B. Alanin, Valin, Leucin, Isoleucin, Prolin, Phenylalanin, Methionin, Tryptophan), beta-verzweigten Seitenketten (z.B. Threonin, Valin, Isoleucin) und aromatischen Seitenketten (z.B. Tyrosin, Phenylalanin, Tryptophan, Histidin). Ein vorhergesagter nicht-essentieller Aminosäurerest in einer Desaturase wird somit vorzugsweise durch einen anderen Aminosäurerest aus der gleichen Seitenkettenfamilie ausgetauscht. Alternativ können bei einer anderen Ausführungsform die Mutationen zufallsgemäß über die gesamte oder einen Teil der Desaturase-kodierenden Sequenz eingebracht werden, z.B. durch Sättigungsmutagenese, und die resultierenden Mutanten können nach der hier beschriebenen Desaturase-Aktivität durchmustert werden, um Mutanten zu identifizieren, die Desaturaseaktivität beibehalten. Nach der Mutagenese einer der Sequenzen der SEQ ID NO: 1, 3, 5 oder 11 kann das kodierte Protein rekombinant exprimiert werden, und die Aktivität des Proteins kann z.B. unter Verwendung der hier beschriebenen Tests (siehe Beispielteil) bestimmt werden.

Zusätzlich zu den Nukleinsäuremolekülen, welche die vorstehend beschriebenen Desaturasen kodieren, betrifft ein weiterer Aspekt der Erfindung isolierte Nukleinsäuremoleküle, die "Antisense" zu den erfundungsgemäßen Nukleinsäuresequenzen sind. Eine "Antisense"-Nukleinsäure umfasst eine Nukleotidsequenz, die zu einer "Sense"-Nukleinsäure, welche ein Protein kodiert, komplementär ist, z.B. komplementär zum kodierenden Strang eines doppelsträngigen cDNA-Moleküls oder komplementär zu einer mRNA-Sequenz.

Eine Antisense-Nukleinsäure kann folglich über Wasserstoffbrückenbindungen an eine Sense-Nukleinsäure binden. Die Antisense-Nukleinsäure kann zu einem gesamten Desaturase-kodierenden Strang oder nur zu einem Teil davon komplementär sein. Bei einer Ausführungsform ist ein Antisense-Nukleinsäuremolekül "Antisense" zu einem "kodierenden Bereich" des kodierenden Strangs einer Nukleotidsequenz, die eine Desaturase kodiert. Der Begriff "kodierender Bereich" betrifft den Bereich der Nukleotidsequenz, der Codons umfasst, die in Aminosäurereste translatiert werden (z.B. den gesamten kodierenden Bereich, der mit dem Stopcodon beginnt und endet, d.h. dem letzten Codon vor dem Stopcodon). Bei einer weiteren Ausführungsform ist das Antisense-Nukleinsäuremolekül "Antisense" zu einem "nicht-kodierenden Bereich" des kodierenden Strangs einer Nukleotidsequenz, die Desaturase kodiert. Der Begriff "nicht-kodierender Bereich" betrifft 5'- und 3'-Sequenzen, die den kodierenden Bereich flankieren und nicht in Aminosäuren translatiert werden (d.h. die man auch als 5'- und 3'-untranslatierte Bereiche bezeichnet).

Unter Voraussetzung der hier offenbarten Desaturase-kodierenden Sequenzen des kodierenden Stranges (z.B. die in SEQ ID NO: 1, 3, 5 oder 11 dargestellten Sequenzen) können erfindungsgemäß Antisense-Nukleinsäuren gemäß den Regeln der Watson-Crick-Basenpaarung gestaltet werden. Das Antisense-Nukleinsäuremolekül kann komplementär zum gesamten kodierenden Bereich von Desaturase-mRNA sein, ist aber stärker bevorzugt ein Oligonukleotid, das nur zu einem Teil des kodierenden oder nicht-kodierenden Bereichs von Desaturase-mRNA "Antisense" ist. Das Antisense-Oligonukleotid kann z.B. zu dem Bereich, der die Translationsstartstelle von Desaturase-mRNA umgibt, komplementär sein. Ein Antisense-Oligonukleotid kann z.B. etwa 5, 10, 15, 20, 25, 30, 35, 40, 45 oder 50 und mehr Nukleotide lang sein. Eine erfindungsgemäß Antisense-Nukleinsäure kann unter Verwendung chemischer Synthese und enzymatischer Ligationsreaktionen mittels im Fachgebiet bekannter Verfahren konstruiert werden. Eine Antisense-Nukleinsäure (z.B. ein Antisense-Oligonukleotid) kann z.B. chemisch synthetisiert werden, wobei natürlich vorkommende Nukleotide oder verschiedentlich modifizierte Nukleotide verwendet werden, die so gestaltet sind, dass sie die biologische Stabilität der Moleküle erhöhen oder die physikalische Stabilität des zwischen der Antisense- und der Sense-Nukleinsäure gebildeten Duplexes erhöhen, beispielsweise können Phosphorthioat-Derivate und acridinsubstituierte Nukleotide verwendet werden. Beispiele für modifizierte Nukleotide, die zur Erzeugung der Antisense-Nukleinsäure verwendet werden können, sind u.a. 5-Fluoruracil, 5-Bromuracil, 5-Chloruracil, 5-Ioduracil, Hypoxanthin, Xanthin, 4-Acetylcytosin, 5-(Carboxyhydroxymethyl)uracil, 5-Carboxymethylaminomethyl-2-

thiouridin, 5-Carboxymethylaminomethyluracil, Dihydouracil,
Beta-D-Galactosylqueosin, Inosin, N6-Isopentenyladenin, 1-Methyl-
guanin, 1-Methylinosin, 2,2-Dimethylguanin, 2-Methyladenin,
2-Methylguanin, 3-Methylcytosin, 5-Methylcytosin, N6-Adenin,
5 7-Methylguanin, 5-Methylaminomethyluracil, 5-Methoxyaminomethyl-
2-thiouracil, Beta-D-Mannosylqueosin, 5'-Methoxycarboxymethyl-
uracil, 5-Methoxyuracil, 2-Methylthio-N6-isopentyladenin, Uracil-
5-oxyessigsäure (v), Wybutoxosin, Desaturaseudouracil, Queosin,
2-Thiocytosin, 5-Methyl-2-thiouracil, 2-Thiouracil, 4-Thiouracil,
10 5-Methyluracil, Uracil-5-oxyessigsäuremethylene, Uracil-5-
oxyessigsäure (v), 5-Methyl-2-thiouracil, 3-(3-Amino-3-N-2-
carboxypropyl)uracil, (acp3)w und 2,6-Diaminopurin. Die Anti-
sense-Nukleinsäure kann alternativ biologisch unter Verwendung
eines Expressionsvektors hergestellt werden, in den eine Nuklein-
15 säure in Antisense-Richtung subkloniert worden ist (d.h. RNA,
die von der eingebrachten Nukleinsäure transkribiert wird, ist
zu einer Zielnukleinsäure von Interesse in Antisense-Richtung
orientiert, was im nachstehenden Unterabschnitt weiter
beschrieben ist).
20 Die erfindungsgemäßen Antisense-Nukleinsäuremoleküle werden
üblicherweise an eine Zelle verabreicht oder in situ erzeugt, so
dass sie mit der zellulären mRNA und/oder der genomischen DNA,
die eine Desaturase kodiert, hybridisieren oder daran binden,
um dadurch die Expression des Proteins, z.B. durch Hemmung der
25 Transkription und/oder Translation, zu hemmen. Die Hybridisierung
kann durch herkömmliche Nukleotidkomplementarität unter Bildung
eines stabilen Duplexes oder z.B. im Fall eines Antisense-
Nukleinsäuremoleküls, das DNA-Duplices bindet, durch spezifische
Wechselwirkungen in der großen Furche der Doppelhelix erfolgen.
30 Das Antisense-Molekül kann so modifiziert sein, dass es spezi-
fisch an einen Rezeptor oder an ein auf einer ausgewählten Zell-
oberfläche exprimierte Antigen bindet, z.B. durch Binden des
Antisense-Nukleinsäuremoleküls an ein Peptid oder einen Anti-
körper, das/der an einen Zelloberflächenrezeptor oder ein Antigen
35 bindet. Das Antisense-Nukleinsäuremolekül kann auch unter Ver-
wendung der hier beschriebenen Vektoren den Zellen zugeführt
werden. Zur Erzielung ausreichender intrazellulärer Konzentra-
tionen der Antisense-Moleküle sind Vektorkonstrukte, in denen
sich das Antisense-Nukleinsäuremolekül unter der Kontrolle eines
40 starken prokaryotischen, viralen oder eukaryotischen, einschließ-
lich pflanzlichen, Promoters befindet, bevorzugt.

Bei einer weiteren Ausführungsform ist das erfindungsgemäße Anti-
sense-Nukleinsäuremolekül ein α -anomeres Nukleinsäuremolekül. Ein
45 α -anomeres Nukleinsäuremolekül bildet spezifische doppelsträngige
Hybride mit komplementärer RNA, wobei die Stränge im Gegensatz zu
gewöhnlichen β -Einheiten parallel zueinander verlaufen. (Gaultier

43

et al. (1987) Nucleic Acids Res. 15:6625-6641). Das Antisense-Nukleinsäuremolekül kann zudem ein 2'-o-Methylribonukleotid (Inoue et al. (1987) Nucleic Acids Res. 15:6131-6148) oder ein chimäres RNA-DNA-Analogon (Inoue et al. (1987) FEBS Lett. 5 215:327-330) umfassen.

Bei einer weiteren Ausführungsform ist eine erfindungsgemäße Antisense-Nukleinsäure ein Ribozym. Ribozyme sind katalytische RNA-Moleküle mit Ribonukleaseaktivität, die eine einzelsträngige 10 Nukleinsäure, wie eine mRNA, spalten können, zu der sie einen komplementären Bereich haben. Somit können Ribozyme (z.B. Hammerhead-Ribozyme (beschrieben in Haselhoff und Gerlach (1988) Nature 334:585-591)) zur katalytischen Spaltung von Desaturase-mRNA-Transkripten verwendet werden, um dadurch die Translation von 15 Desaturase-mRNA zu hemmen. Ein Ribozym mit Spezifität für eine Desaturase-kodierende Nukleinsäure kann auf der Basis der Nukleotidsequenz einer der in SEQ ID NO: 1, 3, 5 oder 11 offenbarten Desaturase-cDNA (d.h. oder auf der Basis einer gemäß den in dieser Erfindung gelehrt Verfahren zu isolierenden heterologen 20 Sequenz gestaltet werden. Beispielsweise kann ein Derivat einer Tetrahymena-L-19-IVS-RNA konstruiert werden, wobei die Nukleotidsequenz der aktiven Stelle komplementär zu der Nukleotidsequenz ist, die in einer Desaturase-kodierenden mRNA gespalten werden soll. Siehe z.B. Cech et al., US-Patent Nr. 4,987,071 und Cech 25 et al., US-Patent Nr. 5,116,742. Alternativ kann Desaturase-mRNA zur Selektion einer katalytischen RNA mit einer spezifischen Ribonukleaseaktivität aus einem Pool von RNA-Molekülen verwendet werden. Siehe z.B. Bartel, D., und Szostak, J.W. (1993) Science 261:1411-1418.

30 Alternativ lässt sich die Desaturase-Gen-Expression hemmen, indem Nukleotidsequenzen, die komplementär zum regulatorischen Bereich einer Desaturase-Nuklectidsequenz (z.B. einem Desaturase-Promotor und/oder -Enhancer) sind, so dirigiert werden, dass 35 Dreifachhelix-Strukturen gebildet werden, welche die Transkription eines Desaturase-Gens in Zielzellen hemmen. Siehe allgemein Helene, C. (1991) Anticancer Drug Res. 6(6) 569-84; Helene, C., et al. (1992) Ann. N. Y. Acad. Sci. 660:27-36; und Maher, L.J. (1992) Bioassays 14(12):807-815.

40 B. Genkonstrukt (= Nukleinsäurekonstrukt, -fragment oder Expressionskassette)

Unter der erfindungsgemäßen Expressionskassette sind die in 45 SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 11 genannten Sequenzen, die sich als Ergebnis des genetischen Codes und/oder deren funktionellen oder nicht funktionellen Derivate

zu verstehen, die mit einem oder mehreren Regulationssignalen vorteilhaft erweise zur Erhöhung der Genexpression funktionell verknüpft wurden und welche vorteilhaft die Expression der codierenden Sequenz in der Wirtszelle steuern. Diese regulatorischen

5 Sequenzen sollen die gezielte Expression der Gene und der Proteinexpression ermöglichen. Dies kann beispielsweise je nach Wirtsorganismus bedeuten, dass das Gen erst nach Induktion exprimiert und/oder überexprimiert wird, oder dass es sofort exprimiert und/oder überexprimiert wird. Beispielsweise handelt

10 es sich bei diesen regulatorischen Sequenzen um Sequenzen an die Induktoren oder Repressoren binden und so die Expression der Nukleinsäure regulieren. Zusätzlich zu diesen neuen Regulationssequenzen oder anstelle dieser Sequenzen kann die natürliche Regulation dieser Sequenzen vor den eigentlichen Strukturgenen

15 noch vorhanden sein und gegebenenfalls genetisch verändert worden sein, so dass die natürliche Regulation ausgeschaltet und die Expression der Gene erhöht wurde. Das Genkonstrukt kann aber auch einfacher aufgebaut sein, das heißt es wurden keine zusätzlichen Regulationssignale vor die Nukleinsäuresequenz oder

20 dessen Derivate inseriert und der natürliche Promotor mit seiner Regulation wurde nicht entfernt. Stattdessen wurde die natürliche Regulationssequenz so mutiert, dass keine Regulation mehr erfolgt und/oder die Genexpression gesteigert wird. Diese veränderten Promotoren können in Form von Teilsequenzen (= Promotor mit

25 Teilen der erfindungsgemäßen Nukleinsäuresequenzen) auch allein vor das natürliche Gen zur Steigerung der Aktivität gebracht werden. Das Genkonstrukt kann außerdem vorteilhaft erweise auch eine oder mehrere sogenannte "enhancer Sequenzen" funktionell verknüpft mit dem Promotor enthalten, die eine erhöhte Expression

30 der Nukleinsäuresequenz ermöglichen. Auch am 3'-Ende der DNA-Sequenzen können zusätzliche vorteilhafte Sequenzen inseriert werden wie weitere regulatorische Elemente oder Terminatoren. Die Δ-5-Desaturase-/Δ-6-Desaturase und/oder Δ-12-Desaturasegene können in einer oder mehreren Kopien in der Expressionskassette

35 (= Genkonstrukt) enthalten sein.

Die regulatorischen Sequenzen bzw. Faktoren können dabei wie oben beschrieben vorzugsweise die Genexpression der eingeführten Gene positiv beeinflussen und dadurch erhöhen. So kann eine Ver-

40 stärkung der regulatorischen Elemente vorteilhaft erweise auf der Transkriptionsebene erfolgen, indem starke Transkriptionssignale wie Promotoren und/oder "Enhancer" verwendet werden. Daneben ist aber auch eine Verstärkung der Translation möglich, indem beispielsweise die Stabilität der mRNA verbessert wird.

Eine weitere Ausführungsform der Erfindung sind ein oder mehrere Genkonstrukte, die eine oder mehrere Sequenzen enthalten, die durch Seq ID NO: 1, 3, 5, 7, 9 oder 11 definiert sind und gem. SEQ ID NO: 2, 4, 6, 8, 10 oder 12 Polypeptide kodieren. Dabei

5 stammen SEQ ID NO: 1, 3, 5, 7 und 11 von Desaturasen während SEQ ID NO: 9 für eine Elongase codiert. Desaturasen codierende Enzyme, die eine Doppelbindung in Δ -5-, Δ -6- oder Δ -12-Position einführen, wobei das Substrat ein, zwei, drei oder vier Doppelbindungen aufweisen. Die in SEQ ID NO: 9 dargestellte Sequenz

10 codiert für eine Enzymaktivität, die eine Fettsäure um mindestens zwei Kohlenstoffatome verlängert sowie ihre Homologen, Derivate oder Analoga, die funktionsfähig mit einem oder mehreren Regulationssignalen, vorteilhafterweise zur Steigerung der Genexpression, verbunden sind. Beispiele für diese Regula-

15 tionssequenzen sind Sequenzen, an die Induktoren oder Repressoren binden und so die Expression der Nukleinsäure regulieren. Zusätzlich zu diesen neuen Regulationssequenzen kann die natürliche Regulation dieser Sequenzen vor den eigentlichen Strukturgenen noch vorhanden sein und, wenn geeignet, genetisch modifiziert

20 werden sein, so dass die natürliche Regulation ausgeschaltet worden ist und die Expression der Gene gesteigert worden ist. Das Genkonstrukt kann jedoch auch eine einfachere Struktur haben, d.h. dass keine zusätzlichen Regulationssignale vor der Sequenz SEQ ID NO: 1, 3, 5 oder 11 oder ihren Homologen inseriert worden

25 sind und der natürliche Promotor mit seiner Regulation nicht deletiert worden ist. Statt dessen ist die natürliche Regulationssequenz so mutiert worden, dass keine Regulation mehr stattfindet und die Genexpression verstärkt ist. Das Genkonstrukt kann außerdem vorteilhafterweise eine oder mehrere sogenannte

30 Enhancer-Sequenzen, die funktionsfähig mit dem Promotor verbunden sind und die gesteigerte Expression der Nukleinsäuresequenz ermöglichen, umfassen. Es ist auch möglich, am 3'-Ende der DNA-Sequenzen zusätzlich vorteilhafte Sequenzen zu inserieren, beispielsweise weitere Regulationselemente oder Terminatoren. Die

35 Desaturasegene und das Elongasegen können im Genkonstrukt in einer oder mehreren Kopien vorliegen. Sie können in einem Genkonstrukt oder mehreren Genkonstrukten vorliegen. Dieses Genkonstrukt oder die Genkonstrukte können zusammen im Wirtszorganismus exprimiert werden. Dabei kann das Genkonstrukt oder

40 die Genkonstrukte in einem oder mehreren Vektoren inseriert sein und frei in der Zelle vorliegen oder aber im Genom inseriert sein. Es ist vorteilhaft für die Insertion weiterer Gene in Organismen, wenn weitere Gene im Genkonstrukt vorliegen.

45 Vorteilhafte Regulationssequenzen für das neue Verfahren liegen beispielsweise in Promotoren vor, wie dem cos-, tac-, trp-, tet-, trp-tet-, lpp-, lac-, lpp-lac-, lacI^Q-, T7-, T5-, T3-, gal-, trc-,

ara-, SP6-, λ - P_R - oder λ - P_L -Promotor und werden vorteilhaft erweise in Gram-negativen Bakterien angewendet. Weitere vorteilhafte Regulationssequenzen liegen beispielsweise in den Gram-positiven Promotoren amy und SPO2, in den Hefe- oder Pilzpromotoren ADC1, 5 Mf α , AC, P-60, CYC1, GAPDH, TEF, rp28, ADH oder in den Pflanzen-prómotoren CaMV/35S [Franck et al., Cell 21 (1980) 285-294], PRP1 [Ward et al., Plant. Mol. Biol. 22 (1993)], SSU, OCS, lib4, usp, STLS1, B33, nos oder im Ubiquitin- oder Phaseolin-Promotor vor. In diesem Zusammenhang vorteilhaft sind ebenfalls induzierbare Promotoren, wie die in EP-A-0 388 186 (Benzylsulfonamid-induzierbar), Plant J. 2, 1992:397-404 (Gatz et al., Tetra-cyclin-induzierbar), EP-A-0 335 528 (Abzisinsäure-induzierbar) oder WO 93/21334 (Ethanol- oder Cyclohexenol-induzierbar) beschriebenen Promotoren. Weitere geeignete Pflanzenpromotoren sind der Promotor von cytosolischer FBPase oder der ST-LSI-Promotor der Kartoffel (Stockhaus et al., EMBO J. 8, 1989, 2445), der Phosphoribosylpyrophosphatamidotransferase-Promotor aus Glycine max (Genbank-Zugangsnr. U87999) oder der in EP-A-0 249 676 beschriebene nodienspezifische Promotor. Besonders vorteilhafte Promotoren sind Promotoren, welche die Expression in Geweben ermöglichen, die an der Fettsäurebiosynthese beteiligt sind. Ganz besonders vorteilhaft sind samenspezifische Promotoren, wie der ausführungsgemäße USP Promotor aber auch andere Promotoren wie der LeB4- (Baeumlein et al., Plant J., 1992, 2 (2) (2):233-239), 25 DC3 (Thomas, Plant Cell 1996, 263:359-368), Phaseolin- oder Napin-Promotor. Weitere besonders vorteilhafte Promotoren sind samenspezifische Promotoren, die für monokotyle oder dikotyle Pflanzen verwendet werden können und in US 5,608,152 (Napin-Promotor aus Raps), WO 98/45461 (Oleosin-Promotor aus 30 Arobidopsis), US 5,504,200 (Phaseolin-Promotor aus Phaseolus vulgaris), WO 91/13980 (Bce4-Promotor aus Brassica), von Baeumlein et al., Plant J., 1992, 2 (2):233-239 (Le34-Promotor aus einer Leguminose) beschrieben sind, wobei sich diese Promotoren für Dikotyledonen eignen. Die folgenden 35 Promotoren eignen sich beispielsweise für Monokotyledonen lpt-2- oder lpt-1-Promotor aus Gerste (WO 95/15389 und WO 95/23230), Hordein-Promotor aus Gerste und andere, in WO 99/16890 beschriebene geeignete Promotoren.

40 Es ist im Prinzip möglich, alle natürlichen Promotoren mit ihren Regulationssequenzen, wie die oben genannten, für das neue Verfahren zu verwenden. Es ist ebenfalls möglich und vorteilhaft, zusätzlich synthetische Promotoren zu verwenden.

45 Das Genkonstrukt kann, wie oben beschrieben, auch weitere Gene umfassen, die in die Organismen eingebracht werden sollen. Es ist möglich und vorteilhaft, in die Wirtsorganismen Regulationsgene,

wie Gene für Induktoren, Repressoren oder Enzyme, welche durch ihre Enzymaktivität in die Regulation eines oder mehrerer Gene eines Biosynthesewegs eingreifen, einzubringen und darin zu exprimieren. Diese Gene können heterologen oder homologen Ursprungs sein. Weiterhin können vorteilhaft im Nukleinsäurekonstrukt bzw. Genkonstrukt weitere Biosynthesegene des Fettsäure- oder Lipidstoffwechsels enthalten sein oder aber diese Gene können auf einem weiteren oder mehreren weiteren Nukleinsäurekonstrukten liegen. Vorteilhaft werden als Biosynthesegene des Fettsäure- oder Lipidstoffwechsels ein Gen ausgewählt aus der Gruppe Acyl-CoA-Dehydrogenase(n), Acyl-ACP[= acyl carrier protein]-Desaturase(n), Acyl-ACP-Thioesterase(n), Fettsäure-Acyl-Transferase(n), Fettsäure-Synthase(n), Fettsäure-Hydroxylase(n), Acetyl-Coenzym A-Carboxylase(n), Acyl-Coenzym A-Oxidase(n), Fettsäure-Desaturase(n), Fettsäure-Acetylensasen, Lipoxygenasen, Triacylglycerol-Lipasen, Allenoxyd-Synthasen, Hydroperoxid-Lyasen oder Fettsäure-Elongase(n) oder deren Kombinationen verwendet.

Genkonstrukte umfassen vorteilhafterweise zur Expression der anderen vorliegenden Gene weitere 3'- und/oder 5'-terminale Regulationssequenzen zur Steigerung der Expression, die in Abhängigkeit vom gewählten Wirtsorganismus und dem Gen oder den Genen für die optimale Expression ausgewählt werden. Diese Regulationssequenzen sollen, wie oben erwähnt, die spezifische Expression der Gene und die Proteinexpression ermöglichen. Dies kann je nach dem Wirtsorganismus beispielsweise bedeuten, dass das Gen nur nach Induktion exprimiert oder überexprimiert wird oder dass es sofort exprimiert und/oder überexprimiert wird.

Die Regulationssequenzen oder -faktoren können außerdem vorzugsweise eine vorteilhafte Wirkung auf die Expression der eingebrachten Gene haben und diese somit steigern. Auf diese Weise ist es möglich, dass die Regulationselemente unter Verwendung starker Transkriptionssignale, wie Promotoren und/oder Enhancer, vorteilhafterweise auf Transkriptionsebene verstärkt werden. Es ist jedoch weiterhin auch möglich, die Translation zum Beispiel durch Verbesserung der mRNA-Stabilität zu verstärken.

40 C. Rekombinante Expressionsvektoren und Wirtszellen

Ein weiterer Aspekt der Erfindung betrifft Vektoren, vorzugsweise Expressionsvektoren, die eine Nukleinsäure enthalten, die eine Desaturase allein (oder einen Teil davon) oder ein unter Punkt b beschriebenes Nukleinsäurekonstrukt in dem die erfindungsgemäße Nukleinsäure allein oder in Kombination mit weiteren Biosynthesegenen des Fettsäure- oder Lipidstoffwechsels

wie Desaturasen oder Elongasen enthalten ist. Wie hier verwendet, betrifft der Begriff "Vektor" ein Nukleinsäuremolekül, das eine andere Nukleinsäure transportieren kann, an welche es gebunden ist. Ein Vektortyp ist ein "Plasmid", was für eine zirkuläre doppelsträngige DNA-Schleife steht, in die zusätzlichen DNA-Segmente ligiert werden können. Ein weiterer Vektortyp ist ein viraler Vektor, wobei zusätzliche DNA-Segmente in das virale Genom ligiert werden können. Bestimmte Vektoren können in einer Wirtszelle, in die sie eingebracht worden sind, autonom replizieren (z.B. Bakterienvektoren mit bakteriellem Replikationsursprung und episomale Säugervektoren). Andere Vektoren (z.B. nicht-episomale Säugervektoren) werden beim Einbringen in die Wirtszelle in das Genom einer Wirtszelle integriert und dadurch zusammen mit dem Wirtsgenom repliziert.

Zudem können bestimmte Vektoren die Expression von Genen, mit denen sie funktionsfähig verbunden sind, steuern. Diese Vektoren werden hier als "Expressionsvektoren" bezeichnet. Gewöhnlich haben Expressionsvektoren, die für DNA-Rekombinationstechniken geeignet sind, die Form von Plasmiden. In der vorliegenden Beschreibung können "Plasmid" und "Vektor" austauschbar verwendet werden, da das Plasmid die am häufigsten verwendete Vektorform ist. Die Erfindung soll jedoch diese anderen Expressionsvektorformen, wie virale Vektoren (z.B. replikationsdefiziente Retroviren, Adenoviren und adenoverwandte Viren), die ähnliche Funktionen ausüben, umfassen. Ferner soll der Begriff Vektor auch andere Vektoren, die dem Fachmann bekannt sind, wie Phagen, Viren, wie SV40, CMV, Baculovirus, Adenovirus, Transposons, IS-Elemente, Phasmide, Phagemide, Cosmide, lineare oder zirkuläre DNA, umfassen.

Die erfindungsgemäßen rekombinanten Expressionsvektoren umfassen eine erfindungsgemäße Nukleinsäure oder ein erfindungsgemäßes Genkonstrukt in einer Form, die sich zur Expression der Nukleinsäure in einer Wirtszelle eignet, was bedeutet, dass die rekombinanten Expressionsvektoren eine oder mehrere Regulationssequenzen, ausgewählt auf der Basis der zur Expression zu verwendenden Wirtszellen, die mit der zu exprimierenden Nukleinsäuresequenz funktionsfähig verbunden ist, umfasst. In einem rekombinanten Expressionsvektor bedeutet "funktionsfähig verbunden", dass die Nukleotidsequenz von Interesse derart an die Regulationssequenz(en) gebunden ist, dass die Expression der Nukleotidsequenz möglich ist und sie aneinander gebunden sind, so dass beide Sequenzen die vorhergesagte, der Sequenz zugeschriebene Funktion erfüllen (z.B. in einem In-vitro-Transkriptions-/Translationssystem oder in einer Wirtszelle, wenn der Vektor in die Wirtszelle eingebracht wird). Der Begriff "Regulationssequenz" soll Promotoren, Enhancer und

andere Expressionskontrollelemente (z.B. Polyadenylierungs-signale) umfassen. Diese Regulationssequenzen sind z.B. beschrieben in Goeddel: Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, CA (1990), oder 5 siehe: Gruber und Crosby, in: Methods in Plant Molecular Biology and Biotechnology, CRC Press, Boca Raton, Florida, Hrsgb.: Glick und Thompson, Kapitel 7, 89-108, einschließlich der Literaturstellen darin. Regulationssequenzen umfassen solche, welche die konstitutive Expression einer Nukleotidsequenz in vielen Wirts-10 zelltypen steuern, und solche, welche die direkte Expression der Nukleotidsequenz nur in bestimmten Wirtszellen unter bestimmten Bedingungen steuern. Der Fachmann weiß, dass die Gestaltung des Expressionsvektors von Faktoren, wie der Auswahl der zu transformierenden Wirtszelle, dem Ausmaß der Expression des 15 gewünschten Proteins usw., abhängen kann. Die erfindungsgemäßen Expressionsvektoren können in Wirtszellen eingebracht werden, um dadurch Proteine oder Peptide, einschließlich Fusionsproteinen oder -peptiden, herzustellen, die von den Nukleinsäuren, wie hier beschrieben, kodiert werden (z.B. Desaturasen, mutante Formen von. 20 Desaturasen, Fusionsproteine usw.).

Die erfindungsgemäßen rekombinannten Expressionsvektoren können zur Expression von Desaturasen und Elongasen in prokaryotischen oder eukaryotischen Zellen gestaltet sein. Beispielsweise können 25 Desaturasegene in bakteriellen Zellen, wie *C. glutamicum*, Insektenzellen (unter Verwendung von Baculovirus-Expressions-vektoren), Hefe- und anderen Pilzzellen (siehe Romanos, M.A., et al. (1992) "Foreign gene expression in yeast: a review", Yeast 8:423-488; van den Hondel, C.A.M.J.J., et al. (1991) 30 "Heterologous gene expression in filamentous fungi", in: More Gene Manipulations in Fungi, J.W. Bennet & L.L. Lasure, Hrsgb., S. 396-428: Academic Press: San Diego; und van den Hondel, C.A.M.J.J., & Punt, P.J. (1991) "Gene transfer systems and vector development for filamentous fungi, in: Applied Molecular Genetics 35 of Fungi, Peberdy, J.F., et al., Hrsgb., S. 1-28, Cambridge University Press: Cambridge), Algen (Falciatore et al., 1999, Marine Biotechnology 1, 3:239-251), Ciliaten der Typen: Holotrichia, Peritrichia, Spirotrichia, Suctoria, Tetrahymena, Paramecium, Colpidium, Glaucoma, Platyophrya, Potomacus, Desaturase-40 udochnilembus, Euplotes, Engelmaniella und Stylonychia, insbesondere der Gattung Stylonychia lemnae, mit Vektoren nach einem Transformationsverfahren, wie beschrieben in WO 98/01572, sowie Zellen vielzelliger Pflanzen (siehe Schmidt, R. und Willmitzer, L. (1988) "High efficiency Agrobacterium tumefaciens-mediated 45 transformation of *Arabidopsis thaliana* leaf and cotyledon explants" Plant Cell Rep.:583-586; Plant Molecular Biology and Biotechnology, C Press, Boca Raton, Florida, Kapitel 6/7,

S. 71-119 (1993); F.F. White, B. Jenes et al., Techniques for Gene Transfer, in: Transgenic Plants, Bd. 1, Engineering and Utilization, Hrsg.: Kung und R. Wu, Academic Press (1993), 128-43; Potrykus, Annu. Rev. Plant Physiol. Plant Molec. Biol. 42 5 (1991), 205-225 (und darin zitierte Literaturstellen) oder Säugerzellen exprimiert werden. Geeignete Wirtszellen werden ferner erörtert in Goeddel, Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, CA (1990). Der rekombinante Expressionsvektor kann alternativ, zum Beispiel 10 unter Verwendung von T7-Promotor-Regulationssequenzen und T7-Polymerase, in vitro transkribiert und translatiert werden.

Die Expression von Proteinen in Prokaryoten erfolgt meist mit Vektoren, die konstitutive oder induzierbare Promotoren enthalten, welche die Expression von Fusions- oder nicht-Fusionsproteinen steuern. Fusionsvektoren fügen eine Reihe von Aminosäuren an ein darin kodiertes Protein an, gewöhnlich am Amino-terminus des rekombinanten Proteins, aber auch am C-Terminus oder fusioniert innerhalb geeigneter Bereiche in den Proteinen. 15 Diese Fusionsvektoren haben gewöhnlich drei Aufgaben: 1) die Verstärkung der Expression von rekombinantem Protein; 2) die Erhöhung der Löslichkeit des rekombinanten Proteins und 3) die Unterstützung der Reinigung des rekombinanten Proteins durch Wirkung als Ligand bei der Affinitätsreinigung. Bei Fusions- 20 Expressionsvektoren wird oft eine proteolytische Spaltstelle an der Verbindungsstelle der Fusionseinheit und des rekombinanten Proteins eingebracht, so dass die Abtrennung des rekombinanten Proteins von der Fusionseinheit nach der Reinigung des Fusionsproteins möglich ist. Diese Enzyme und ihre entsprechenden 25 30 Erkennungssequenzen umfassen Faktor Xa, Thrombin und Enterokinase.

Typische Fusions-Expressionsvektoren sind u.a. pGEX (Pharmacia Biotech Inc; Smith, D.B., und Johnson, K.S. (1988) Gene 35 67:31-40), pMAL (New England Biolabs, Beverly, MA) und pRITS (Pharmacia, Piscataway, NJ), bei denen Glutathion-S-Transferase (GST), Maltose E-bindendes Protein bzw. Protein A an das rekombinante Zielprotein fusioniert wird. Bei einer Ausführungsform ist die Desaturase-kodierende Sequenz in einen pGEX- 40 Expressionsvektor kloniert, so dass ein Vektor erzeugt wird, der ein Fusionsprotein kodiert, das vom N-Terminus zum C-Terminus GST-Thrombin-Spaltstelle-X-Protein umfasst. Das Fusionsprotein kann durch Affinitätschromatographie unter Verwendung von Glutathion-Agarose-Harz gereinigt werden. Rekombinante Desaturase, die 45 nicht an GST fusioniert ist, kann durch Spaltung des Fusionsproteins mit Thrombin gewonnen werden.

Beispiele für geeignete induzierbare nicht-Fusions-E. coli-Expressionsvektoren sind u.a. pTrc (Amann et al. (1988) Gene 69:301-315) und pET 11d (Studier et al., Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, Kalifornien (1990) 60-89). Die Zielgenexpression vom pTrc-Vektor beruht auf der Transkription durch Wirts-RNA-Polymerase von einem Hybrid-trp-lac-Fusionspromotor. Die Zielgenexpression aus dem pET 11d-Vektor beruht auf der Transkription von einem T7-gn10-lac-Fusions-Promotor, die von einer coexprimierten viralen RNA-Polymerase (T7 gn1) vermittelt wird. Diese virale Polymerase wird von den Wirtsstämmen BL21 (DE3) oder HMS174 (DE3) von einem residenten λ -Prophagen bereitgestellt, der ein T7 gn1-Gen unter der Transkriptionskontrolle des lacUV 5-Promotors birgt.

15 Andere in prokaryotischen Organismen geeignete Vektoren sind dem Fachmann bekannt, diese Vektoren sind beispielsweise in E. coli pLG338, pACYC184, die pBR-Reihe, wie pBR322, die pUC-Reihe, wie pUC18 oder pUC19, die M113mp-Reihe, pKC30, pRep4, pHs1, pHs2, pPLc236, pMBL24, pLG200, pUR290, pIN-III¹¹³-B1, λ gt11 or pBdCI, 20 in Streptomyces pIJ101, pIJ364, pIJ702 oder pIJ361, in Bacillus pUB110, pC194 oder pBD214, in Corynebacterium pSA77 oder pAJ667. Eine Strategie zur Maximierung der Expression von rekombinantem Protein ist die Expression des Proteins in einem Wirtsbakterium, dessen Fähigkeit zur proteolytischen Spaltung des rekombinanten 25 Proteins gestört ist (Gottesman, S., Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, Kalifornien (1990) 119-128). Eine weitere Strategie ist die Veränderung der Nukleinsäuresequenz der in einen Expressionsvektor zu inserierenden Nukleinsäure, so dass die einzelnen Codons für jede Amino- 30 säure diejenigen sind, die vorzugsweise in einem zur Expression ausgewählten Bakterium, wie C. glutamicum, verwendet werden (Wada et al. (1992) Nucleic Acids Res. 20:2111-2118). Diese Veränderung der erfindungsgemäßen Nukleinsäuresequenzen erfolgt durch Standard-DNA-Synthesetechniken.

35 Bei einer weiteren Ausführungsform ist der Desaturase-Expressionsvektor ein Hefe-Expressionsvektor. Beispiele für Vektoren zur Expression in der Hefe S. cerevisiae umfassen pYeDesaturasec1 (Baldari et al. (1987) Embo J. 6:229-234), 40 pMFA (Kurjan und Herskowitz (1982) Cell 30:933-943), pJRY88 (Schultz et al. (1987) Gene 54:113-123) sowie pYES2 (Invitrogen Corporation, San Diego, CA). Vektoren und Verfahren zur Konstruktion von Vektoren, die sich zur Verwendung in anderen Pilzen, wie den filamentösen Pilzen, eignen, umfassen diejenigen, 45 die eingehend beschrieben sind in: van den Hondel, C.A.M.J.J., & Punt, P.J. (1991) "Gene transfer systems and vector development for filamentous fungi, in: Applied Molecular Genetics of fungi,

52

J.F. Pebernry et al., Hrsgb., S. 1-28, Cambridge University Press: Cambridge, oder in: More Gene Manipulations in Fungi [J.W. Bennet & L.L. Lasure, Hrsgb., S. 396-428: Academic Press: San Diego]. Weitere geeignete Hefevektoren sind beispielsweise pAG-1, YEp6,
5 YEp13 oder pEMBLYe23.

Alternativ können die erfindungsgemäßen Desaturasen in Insektenzellen unter Verwendung von Baculovirus-Expressionsvektoren exprimiert werden. Baculovirus-Vektoren, die zur Expression von 10 Proteinen in gezüchteten Insektenzellen (z.B. Sf9-Zellen) verfügbar sind, umfassen die pAc-Reihe (Smith et al. (1983) Mol. Cell Biol. 3:2156-2165) und die pVL-Reihe (Lucklow und Summers (1989) Virology 170:31-39).

15 Die oben genannten Vektoren bieten nur einen kleinen Überblick über mögliche geeignete Vektoren. Weitere Plasmide sind dem Fachmann bekannt und sind zum Beispiel beschrieben in: Cloning Vectors (Hrsgb. Pouwels, P.H., et al., Elsevier, Amsterdam-New York-Oxford, 1985, ISBN 0 444 904018).

20 Bei noch einer weiteren Ausführungsform wird eine erfindungsgemäße Nukleinsäure in Säugerzellen unter Verwendung eines Säuger-Expressionsvektors exprimiert. Unter Säugern werden im Sinne der Erfindung alle nicht-humanen Säuger verstanden.

25 Beispiele für Säuger-Expressionsvektoren umfassen pCDM8 (Seed, B. (1987) Nature 329:840) und pMT2PC (Kaufman et al. (1987) EMBO J. 6:187-195). Bei der Verwendung in Säugerzellen werden die Kontrollfunktionen des Expressionsvektors oft von viralen Regulationselementen bereitgestellt. Üblicherweise verwendete 30 Promotoren stammen z.B. aus Polyoma, Adenovirus2, Cytomegalievirus und Simian Virus 40. Weitere geeignete Expressionssysteme für prokaryotische und eukaryotische Zellen siehe in den Kapiteln 16 und 17 von Sambrook, J., Fritsch, E.F., und Maniatis, T., Molecular Cloning: A Laboratory Manual, 2. Auflage, Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

Bei einer anderen Ausführungsform kann der rekombinante Säuger-Expressionsvektor die Expression der Nukleinsäure vorzugsweise in 40 einem bestimmten Zelltyp steuern (z.B. werden gewebespezifische Regulationselemente zur Expression der Nukleinsäure verwendet). Gewebespezifische Regulationselemente sind im Fachgebiet bekannt. Nicht beschränkende Beispiele für geeignete gewebespezifische Promotoren sind u.a. der Albuminpromotor (leberspezifisch; 45 Pinkert et al. (1987) Genes Dev. 1:268-277), lymphoidspezifische Promotoren (Calame und Eaton (1988) Adv. Immunol. 43:235-275), insbesondere Promotoren von T-Zellrezeptoren (Winoto und

Baltimore (1989) EMBO J. 8:729-733) und Immunglobulinen (Banerji et al. (1983) Cell 33:729-740; Queen und Baltimore (1983) Cell 33:741-748), neuronspezifische Promotoren (z.B. Neurofilament-Promotor; Byrne und Ruddle (1989) PNAS 86:5473-5477), pankreas-
5 spezifische Promotoren (Edlund et al., (1985) Science 230:912-916) und milchdrüsenspezifische Promotoren (z.B. Milch-serum-Promotor; US-Patent Nr. 4,873,316 und Europäische Patent-anmeldung-Veröffentlichung Nr. 264,166). Auch entwicklungs-regulierte Promotoren sind umfasst, z.B. die hox-Promotoren
10 der Maus (Kessel und Gruss (1990) Science 249:374-379) und der Fetoprotein-Promotor (Campes und Tilghman (1989) Genes Dev. 3:537-546).

Bei einer weiteren Ausführungsform können die erfindungsgemäßen
15 Desaturasen in einzelligen Pflanzenzellen (wie Algen), siehe Falciatore et al., 1999, Marine Biotechnology 1 (3):239-251 und darin zitierte Literaturangaben, und Pflanzenzellen aus höheren Pflanzen (z.B. Spermatophyten, wie Feldfrüchten) exprimiert werden. Beispiele für Pflanzen-Expressionsvektoren
20 umfassen solche, die eingehend beschrieben sind in: Becker, D., Kemper, E., Schell, J., und Masterson, R. (1992) "New plant binary vectors with selectable markers located proximal to the left border", Plant Mol. Biol. 20:1195-1197; und Bevan, M.W. (1984) "Binary Agrobacterium vectors for plant transformation", Nucl. Acids Res. 12:8711-8721; Vectors for Gene Transfer in Higher Plants; in: Transgenic Plants, Bd. 1, Engineering and Utilization, Hrsg.: Kung und R. Wu, Academic Press, 1993, S. 15-38.
25
30 Eine Pflanzen-Expressionskassette enthält vorzugsweise Regulationssequenzen, welche die Genexpression in Pflanzenzellen steuern können und funktionsfähig verbunden sind, so dass jede Sequenz ihre Funktion, wie Termination der Transkription, erfüllen kann, beispielsweise Polyadenylierungssignale. Bevorzugte Polyadenylierungssignale sind diejenigen, die aus Agrobacterium tumefaciens-t-DNA stammen, wie das als Octopinsynthase bekannte Gen 3 des Ti-Plasmids pTiACH5 (Gielen et al., EMBO J. 3 (1984) 835ff.) oder funktionelle Äquivalente davon, aber auch alle anderen in Pflanzen funktionell aktiven Terminatoren
35
40 sind geeignet.

Da die Pflanzengenexpression sehr oft nicht auf Transkriptions-ebenen beschränkt ist, enthält eine Pflanzen-Expressionskassette vorzugsweise andere funktionsfähig verbunden Sequenzen, wie Translationsenhancer, beispielsweise die Overdrive-Sequenz, welche die 5'-untranslatierte Leader-Sequenz aus Tabakmosaik-
45

virus, die das Protein/RNA-Verhältnis erhöht, enthält (Gallie et al., 1987, Nucl. Acids Research 15:8693-8711).

Die Pflanzengenexpression muss funktionsfähig mit einem geeigneten Promotor verbunden sein, der die Genexpression auf rechtzeitige, zell- oder gewebespezifische Weise durchführt. Bevorzugt sind Promotoren, welche die konstitutive Expression herbeiführen (Benfey et al., EMBO J. 8 (1989) 2195-2202), wie diejenigen, die von Pflanzenviren stammen, wie 35S CaMV (Franck et al., Cell 21 (1980) 285-294), 19S CaMV (siehe auch US 5352605 und WO 84/02913) oder Pflanzenpromotoren, wie der in US 4,962,028 beschriebene der kleinen Untereinheit der Rubisco.

Andere bevorzugte Sequenzen für die Verwendung zur funktionsfähigen Verbindung in Pflanzengenexpressions-Kassetten sind Targeting-Sequenzen, die zur Steuerung des Genproduktes in sein entsprechendes Zellkompartiment notwendig sind (siehe eine Übersicht in Kermode, Crit. Rev. Plant Sci. 15, 4 (1996) 285-423 und darin zitierte Literaturstellen), beispielsweise in die Vakuole, den Zellkern, alle Arten von Plastiden, wie Amyloplasten, Chloroplasten, Chromoplasten, den extrazellulären Raum, die Mitochondrien, das Endoplasmatische Retikulum, Ölkörper, Peroxisomen und andere Kompartimente von Pflanzenzellen.

Die Pflanzengenexpression lässt sich auch über einen chemisch induzierbaren Promotor erleichtern (siehe eine Übersicht in Gatz 1997, Annu. Rev. Plant Physiol. Plant Mol. Biol., 48:89-108). Chemisch induzierbare Promotoren eignen sich besonders, wenn gewünscht wird, dass die Genexpression auf zeitspezifische Weise erfolgt. Beispiele für solche Promotoren sind ein Salicylsäure-induzierbarer Promotor (WO 95/19443), ein Tetracyclin-induzierbarer Promotor (Gatz et al. (1992) Plant J. 2, 397-404) und ein Ethanol-induzierbarer Promotor.

Auch Promotoren, die auf biotische oder abiotische Stressbedingungen reagieren, sind geeignete Promotoren, beispielsweise der pathogeninduzierte PRP1-Gen-Promotor (Ward et al., Plant. Mol. Biol. 22 (1993) 361-366), der hitzeinduzierbare hsp80-Promotor aus Tomate (US 5,187,267), der kälteinduzierbare Alpha-amylase-Promotor aus Kartoffel (WO 96/12814) oder der durch Wunden induzierbare pinII-Promotor (EP-A-0 375 091).

Es sind insbesondere solche Promotoren bevorzugt, welche die Genexpression in Geweben und Organen herbeiführen, in denen die Lipid- und Ölbiosynthese stattfindet, in Samenzellen, wie den Zellen des Endosperms und des sich entwickelnden Embryos. Geeignete Promotoren sind der Napingen-Promotor aus Raps

(US 5,608,152), der USP-Promotor aus *Vicia faba* (Baeumlein et al., Mol Gen Genet, 1991, 225 (3):459-67), der Oleosin-Promotor aus *Arabidopsis* (WO 98/45461), der Phaseolin-Promotor aus *Phaseolus vulgaris* (US 5,504,200), der Bce4-Promotor aus 5 *Brassica* (WO 91/13980) oder der Legumin-B4-Promotor (LeB4; Baeumlein et al., 1992, Plant Journal, 2 (2):233-9) sowie Promotoren, welche die samenspezifische Expression in Monokotyledonen-Pflanzen, wie Mais, Gerste, Weizen, Roggen, Reis usw. herbeiführen. Geeignete beachtenswerte Promotoren sind der lpt2- 10 oder lpt1-Gen-Promotor aus Gerste (WO 95/15389 und WO 95/23230) oder die in WO 99/16890 beschriebenen (Promotoren aus dem Gersten-Hordein-Gen, dem Reis-Glutelin-Gen, dem Reis-Oryzin-Gen, dem Reis-Prolamín-Gen, dem Weizen-Gliadin-Gen, Weizen-Glutelin-Gen, dem Mais-Zein-Gen, dem Hafer-Glutelin-Gen, dem Sorghum- 15 Kasirin-Gen, dem Roggen-Secalin-Gen).

Insbesondere kann die multiparallele Expression von erfindungsgemäßen Desaturasen allein oder in Kombination mit anderen desaturasen oder Elongasen gewünscht sein. Die Einführung solcher 20 Expressionskassetten kann über eine simultane Transformation mehrerer einzelner Expressionskonstrukte erfolgen oder durch Kombination mehrerer Expressionskassetten auf einem Konstrukt. Auch können mehrere Vektoren mit jeweils mehreren Expressionskassetten transformiert und auf die Wirtszelle 25 übertragen werden.

Ebenfalls besonders geeignet sind Promotoren, welche die plastidenspezifische Expression herbeiführen, da Plastiden das Kompartiment sind, in dem die Vorläufer sowie einige Endprodukte der Lipidbiosynthese synthetisiert werden. Geeignete 30 Promotoren, wie der virale RNA-Polymerase-Promotor, sind beschrieben in WO 95/16783 und WO 97/06250, und der clpP-Promotor aus *Arabidopsis*, beschrieben in WO 99/46394.

35 Die Erfindung stellt zudem einen rekombinanten Expressionsvektor bereit, umfassend ein erfindungsgemäßes DNA Molekül, das in Antisense-Richtung in den Expressionsvektor kloniert ist. d.h. das DNA-Molekül ist derart mit einer regulatorischen Sequenz funktionsfähig verbunden, dass die Expression (durch 40 Transkription des DNA-Moleküls) eines RNA-Moleküls, das zur Desaturase-mRNA "Antisense" ist, ermöglicht wird. Es können Regulationssequenzen ausgewählt werden, die funktionsfähig mit einer in Antisense-Richtung klonierten Nukleinsäure verbunden sind und die kontinuierliche Expression des Antisense-RNA-Moleküls in einer Vielzahl von Zelltypen steuern, zum Beispiel können 45 virale Promotoren und/oder Enhancer oder Regulationssequenzen ausgewählt werden, welche die konstitutive, gewebespezifische

oder zelltypspezifische Expression von Antisense-RNA steuern. Der Antisense-Expressionsvektor kann in Form eines rekombinanten Plasmids, Phagemids oder attenuierten Virus vorliegen, in dem Antisense-Nukleinsäuren unter der Kontrolle eines hochwirksamen 5 regulatorischen Bereichs produziert werden, dessen Aktivität durch den Zelltyp bestimmt werden kann, in den der Vektor eingebracht worden ist. Eine Erläuterung der Regulation der Genexpression mittels Antisense-Genen siehe in Weintraub, H., et al., Antisense-RNA as a molecular tool for genetic analysis, 10 Reviews - Trends in Genetics, Bd. 1(1) 1986.

Ein weiterer Aspekt der Erfindung betrifft Wirtszellen, in die ein erfindungsgemäßer rekombinanter Expressionsvektor eingebracht worden ist. Die Begriffe "Wirtszelle" und "rekombinante Wirts- 15 zelle" werden hier untereinander austauschbar verwendet. Selbstverständlich betreffen diese Begriffe nicht nur die bestimmte Zielzelle, sondern auch die Nachkommen oder potentiellen Nachkommen dieser Zelle. Da in aufeinanderfolgenden Generationen aufgrund von Mutation oder Umwelteinflüssen bestimmte Modifikationen 20 auftreten können, sind diese Nachkommen nicht unbedingt mit der Parentalzelle identisch, sind jedoch immer noch vom Umfang des Begriffs, wie hier verwendet, umfasst.

Unter Rekombinant oder Transgen beispielsweise rekombinannten 25 Expressionsvektor oder rekombinanten Wirt oder Wirtszellen im Sinne der Erfindung ist zu verstehen, dass die erfindungsgemäßen Nukleinsäuren und/oder deren natürliche Regulationssequenzen an 5' und 3'-Position der Nukleinsäuren nicht in ihrer natürlichen Umgebung sind, das heißt entweder wurde die Lage der Sequenzen im 30 Herkunftorganismus verändert oder in diesem wurden die Nuklein-säuresequenzen und/oder die Regulationssequenzen mutiert oder die erfindungsgemäßen Nukleinsäuresequenzen wurden in einen anderen Organismus als den Herkunftorganismus verbracht oder deren Regulationssequenzen. Auch Kombinationen dieser Veränderungen 35 sind möglich. Unter natürlicher Umgebung ist die Lage einer Nukleinsäuresequenz in einem Organismus zu verstehen, wie er in der Natur vorkommt.

Eine Wirtszelle kann eine prokaryotische oder eukaryotische 40 Zelle sein. Zum Beispiel kann eine Desaturase in Bakterienzellen; wie C. glutamicum, Insektenzellen, Pilzzellen oder Sägerzellen (wie Chinesischer Hamster-Ovarzellen (CHO) oder COS-Zellen), Algen, Ciliaten, Pflanzenzellen, Pilzen oder anderen Mikro-organismen, wie C. glutamicum, exprimiert werden. Andere 45 geeignete Wirtszellen sind dem Fachmann geläufig.

Vektor-DNA lässt sich in prokaryotische oder eukaryotische Zellen über herkömmliche Transformations- oder Transfektionstechniken einbringen. Die Begriffe "Transformation" und "Transfektion", Konjugation und Transduktion, wie hier verwendet, sollen eine 5 Vielzahl von im Stand der Technik bekannten Verfahren zum Einbringen fremder Nukleinsäure (z.B. DNA) in eine Wirtszelle, einschließlich Calciumphosphat- oder Calciumchlorid-Copräzipitation, DEAE-Dextran-vermittelte Transfektion, Lipofektion, natürliche Kompetenz, chemisch vermittelter Transfer, Elektroporation oder Teilchenbeschuss, umfassen. Geeignete Verfahren 10 zur Transformation oder Transfektion von Wirtszellen, einschließlich Pflanzenzellen, lassen sich finden in Sambrook et al. (Molecular Cloning: A Laboratory Manual., 2. Aufl., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold 15 Spring Harbor, NY, 1989) und anderen Labor-Handbüchern, wie Methods in Molecular Biology, 1995, Bd. 44, Agrobacterium protocols, Hrsgb: Gartland und Davey, Humana Press, Totowa, New Jersey.

20 Über die stabile Transfektion von Säugerzellen ist bekannt, dass je nach verwendetem Expressionsvektor und verwendeter Transfektionstechnik nur ein kleiner Teil der Zellen die fremde DNA in ihr Genom integriert. Zur Identifikation und Selektion dieser Integranten wird gewöhnlich ein Gen, das einen selektierbaren Marker (z.B. Resistenz gegen Antibiotika) kodiert, zusammen mit dem Gen von Interesse in die Wirtszellen eingebracht. Bevorzugte selektierbare Marker umfassen solche, welche Resistenz 25 gegen Medikamente, wie G418, Hygromycin und Methotrexat, verleihen, oder in Pflanzen solche, welche Resistenz gegen ein 30 Herbizid, wie Glyphosphat oder Glufosinat, verleihen. Weitere geeignete Marker sind beispielsweise Marker, welche Gene kodieren, die an Biosynthesewegen von zum Beispiel Zuckern oder Aminosäuren beteiligt sind, wie β -Galactosidase, ura3 oder ilv2. Marker, welche Gene, wie Luziferase, gfp oder andere Fluoreszenzgene 35 kodieren, sind ebenfalls geeignet. Diese Marker lassen sich in Mutanten verwenden, in denen diese Gene nicht funktionell sind, da sie beispielsweise mittels herkömmlicher Verfahren deletiert worden sind. Ferner können Marker, welche eine Nukleinsäure 40 kodieren, die einen selektierbaren Marker kodiert, in eine Wirtszelle auf dem gleichen Vektor eingebracht werden, wie derjenige, der eine Desaturase kodiert, oder können auf einem gesonderten Vektor eingebracht werden. Zellen, die mit der eingebrachten Nukleinsäure stabil transfiziert worden sind, können zum Beispiel durch Medikamentenselektion identifiziert werden (z.B. überleben 45 Zellen, die den selektierbaren Marker integriert haben, wohingegen die anderen Zellen absterben).

Zur Erzeugung eines homolog rekombinierten Mikroorganismus wird ein Vektor hergestellt, der zumindest einen Abschnitt eines Desaturasegens enthält, in den eine Deletion, Addition oder Substitution eingebracht worden ist, um dadurch das Desaturasegen 5 zu verändern, z.B. funktionell zu disruptieren. Dieses Desaturasegen ist vorzugsweise ein *Phaeodactylum tricornutum* Desaturasegen, es kann jedoch ein Homologon oder Analogon aus anderen Organismen, sogar aus einer Säuger-, Pilz- oder Insektenquelle verwendet werden. Bei einer bevorzugten Ausführungsform ist der 10 Vektor so gestaltet, dass das endogene Desaturasegen bei homologer Rekombination funktionell disruptiert wird (d.h. nicht länger ein funktionelles Protein kodiert, auch als Knock-out-Vektor bezeichnet). Alternativ kann der Vektor so gestaltet sein, dass das endogene Desaturasegen bei homologer Rekombination 15 mutiert oder anderweitig verändert wird, aber immer noch ein funktionelles Protein kodiert (z.B. kann der stromaufwärts gelegene regulatorische Bereich so verändert sein, dass dadurch die Expression der endogenen Desaturase verändert wird). Zur Erzeugung einer Punktmutation über homologe Rekombination 20 können auch als Chimeroplasty bekannte DNA-RNA-Hybrid verwendete werden, die aus Cole-Strauss et al., 1999, Nucleic Acids Research 27(5):1323-1330 und Kmiec, Gene therapy, 1999, American Scientist, 87(3):240-247 bekannt sind.

25 Im Vektor für die homologe Rekombination ist der veränderte Abschnitt des Desaturasegens an seinem 5'- und 3'-Ende von zusätzlicher Nukleinsäure des Desaturasegens flankiert, so dass homologe Rekombination zwischen dem exogenen Desaturasegen, das auf dem Vektor vorliegt, und einem endogenen Desaturasegen in 30 einem Mikroorganismus oder einer Pflanze möglich ist. Die zusätzliche flankierende Desaturase-Nukleinsäure ist für eine erfolgreiche homologe Rekombination mit dem endogenen Gen hinreichend lang. Gewöhnlich sind im Vektor mehrere hundert Basenpaare bis zu Kilobasen flankierende DNA (sowohl am 5'- als auch am 3'-Ende) 35 enthalten (eine Beschreibung von Vektoren zur homologen Rekombination siehe z.B. in Thomas, K.R., und Capecchi, M.R. (1987) Cell 51:503 oder der Rekombination in *Physcomitrella patens* auf cDNA-Basis in Strepp et al., 1998, Proc. Natl. Acad. Sci. USA 95 (8):4368-4373). Der Vektor wird in einen Mikroorganismus 40 oder eine Pflanzenzelle (z.B. mittels Polyethylenglycol-vermittelte DNA) eingebracht; und Zellen, in denen das eingebrachte Desaturasegen mit dem endogenen Desaturasegen homolog rekombiniert ist, werden unter Verwendung im Fachgebiet bekannter Techniken selektiert.

Bei einer anderen Ausführungsform können rekombinante Organismen, wie Mikroorganismen, hergestellt werden, die ausgewählte Systeme enthalten, welche eine regulierte Expression des eingebrachten Gens ermöglichen. Der Einschluß eines Desaturasegens in einem 5 Vektor, wobei es unter die Kontrolle des lac-Operons gebracht wird, ermöglicht z.B. die Expression des Desaturasegens nur in Gegenwart von IPTG. Diese Regulationssysteme sind im Fachgebiet bekannt.

- 10 Eine erfindungsgemäße Wirtszelle, wie eine prokaryotische oder eukaryotische Wirtszelle, in Kultur oder auf einem Feld wachsend, kann zur Produktion (d.h. Expression) einer Desaturase verwendet werden. In Pflanzen kann zusätzlich ein alternatives Verfahren durch direkten Transfer von DNA in sich entwickelnde Blüten über
- 15 Elektroporation oder Gentransfer mittels Agrobacterium angewendet werden. Die Erfindung stellt folglich ferner Verfahren zur Produktion von Desaturasen unter Verwendung der erfindungsgemäßen Wirtszellen bereit. Bei einer Ausführungsform umfasst das Verfahren die Anzucht der erfindungsgemäßen Wirtszelle (in die
- 20 ein rekombinanter Expressionsvektor, der eine Desaturase kodiert, eingebracht worden ist, oder in deren Genom ein Gen eingebracht worden ist, das eine Wildtyp- oder veränderte Desaturase kodiert) in einem geeigneten Medium, bis die Desaturase produziert worden ist. Das Verfahren umfasst bei einer weiteren Ausführungsform das
- 25 Isolieren der Desaturasen aus dem Medium oder der Wirtszelle.

Wirtszellen, die im Prinzip zum Aufnehmen der erfindungsgemäßen Nukleinsäure, des erfindungsgemäßen Genproduktes oder des erfindungsgemäßen Vektors geeignet sind, sind alle 30 prokaryotischen oder eukaryotischen Organismen. Die vorteilhafterweise verwendeten Wirtsorganismen sind Organismen, wie Bakterien, Pilze, Hefen, Tier- oder Pflanzenzellen. Weitere vorteilhafte Organismen sind Tiere oder vorzugsweise Pflanzen oder Teile davon. Pilze, Hefen oder Pflanzen werden vorzugsweise verwendet, besonders bevorzugt Pilze oder Pflanzen, ganz besonders bevorzugt Pflanzen, wie Ölfruchtpflanzen, die große Mengen an Lipidverbindungen enthalten, wie Raps, Nachtkerze, Canola, Erdnuss, Lein, Soja, Safflor, Sonnenblume, Borretsch, oder Pflanzen, wie Mais, Weizen, Roggen, Hafer, Triticale, Reis, 35 Gerste, Baumwolle, Maniok, Pfeffer, Tagetes, Solanaceen-Pflanzen, wie Kartoffel, Tabak, Aubergine und Tomate, Vicia-Arten, Erbse, Alfalfa, Buschpflanzen (Kaffee, Kakao, Tee), Salix-Arten, Bäume (Ölpalme, Kokosnuss) sowie ausdauernde Gräser und Futterfeldfrüchte. Besonders bevorzugte erfindungsgemäße Pflanzen sind 40 Ölfruchtpflanzen, wie Soja, Erdnuss, Raps, Canola, Lein, Nachtkerze, Sonnenblume, Safflor, Bäume (Ölpalme, Kokosnuss).

D. Isolierte Desaturase

Ein weiterer Aspekt der Erfindung betrifft isolierte Desaturasen und biologisch aktive Teile davon. Ein "isoliertes" oder "ge-reinigtes" Protein oder ein biologisch aktiver Teil davon ist im wesentlichen frei von zellulärem Material, wenn es durch DNA-Rekombinationstechniken produziert wird, oder von chemischen Vorstufen oder andern Chemikalien, wenn es chemisch synthetisiert wird. Der Begriff "im wesentlichen frei von zellulärem Material" umfasst Desaturase-Präparationen, in denen das Protein von zellulären Komponenten der Zellen, in denen es natürlich oder rekombinant produziert wird, getrennt ist. Bei einer Ausführungsform umfasst der Ausdruck "im wesentlichen frei von zellulärem Material" Desaturase-Präparationen mit weniger als etwa 30 % (bezogen auf das Trockengewicht) nicht-Desaturase (hier auch als "verunreinigendes Protein" bezeichnet), stärker bevorzugt weniger als etwa 20 % nicht-Desaturase, noch stärker bevorzugt weniger als etwa 10 % nicht-Desaturase und am stärksten bevorzugt weniger als etwa 5 % nicht-Desaturase. Wenn die Desaturase oder ein biologisch aktiver Teil davon rekombinant hergestellt worden ist, ist sie/er auch im wesentlichen frei von Kulturmedium, d.h. das Kulturmedium macht weniger als etwa 20 %, stärker bevorzugt weniger als etwa 10 % und am stärksten bevorzugt weniger als etwa 5 % des Volumens der Proteinpräparation aus. Der Begriff "im wesentlichen frei von chemischen Vorstufen oder anderen Chemikalien" umfasst Desaturase-Präparationen, in denen das Protein von chemischen Vorstufen oder anderen Chemikalien getrennt ist, die an der Synthese des Proteins beteiligt sind. Bei einer Ausführungsform umfasst der Begriff "im wesentlichen frei von chemischen Vorstufen oder anderen Chemikalien" Desaturase-Präparationen mit weniger als etwa 30 % (bezogen auf das Trockengewicht) chemischen Vorstufen oder nicht-Desaturase-Chemikalien, stärker bevorzugt weniger als etwa 20 % chemischen Vorstufen oder nicht-Desaturase-Chemikalien, noch stärker bevorzugt weniger als etwa 10 % chemischen Vorstufen oder nicht-Desaturase-Chemikalien und am stärksten bevorzugt weniger als etwa 5 % chemischen Vorstufen oder nicht-Desaturase-Chemikalien. Bei bevorzugten Ausführungsformen weisen isolierte Proteine oder biologisch aktive Teile davon keine verunreinigenden Proteine aus dem gleichen Organismus auf, aus dem die Desaturase stammt. Diese Proteine werden gewöhnlich durch rekombinante Expression zum Beispiel Phaeodactylum tricornutum-Desaturase in Pflanzen wie Physcomitrella patens bzw. o.g. oder Mikroorganismen, beispielsweise Bakterien, wie E. coli, Bacillus subtilis, C. glutamicum, Pilzen, wie Mortierella, Hefe, wie Saccharomyces, oder Ciliaten wie Colpidium oder Algen wie Phaeodactylum hergestellt.

Eine erfindungsgemäße isolierte Desaturase oder ein Teil davon kann auch am Stoffwechsel von zum Aufbau von Zellmembranen in Phaeodactylum tricornutum notwendigen Verbindungen oder am Transport von Molekülen über diese Membranen teilnehmen. Bei 5 bevorzugten Ausführungsformen umfasst das Protein oder der Teil davon eine Aminosäuresequenz, die ausreichend homolog zu einer Aminosäuresequenz der SEQ ID NO: 2, 4, 6 oder 12 ist, dass das Protein oder der Teil davon die Fähigkeit, am Stoffwechsel von zum Aufbau von Zellmembranen in Phaeodactylum tricornutum notwendigen Verbindungen oder am Transport von Molekülen über diese Membranen teilzunehmen, beibehält. Der Teil des Proteins ist vorzugsweise ein biologisch aktiver Teil, wie hier beschrieben. Bei einer weiteren bevorzugten Ausführungsform hat eine erfindungsgemäße Desaturase eine der in SEQ ID NO: 2, 4, 6 oder 12 10 gezeigten Aminosäuresequenzen. Bei einer weiteren bevorzugten Ausführungsform hat die Desaturase eine Aminosäuresequenz, die von einer Nukleotidsequenz kodiert wird, die, zum Beispiel unter stringenten Bedingungen, an eine Nukleotidsequenz der SEQ ID NO: 1, 3, 5 oder 11 hybridisiert. Bei noch einer weiteren 15 bevorzugten Ausführungsform hat die Desaturase eine Aminosäuresequenz, die von einer Nukleotidsequenz kodiert wird, die mindestens etwa 50 bis 60 %, vorzugsweise mindestens etwa 60 bis 70 %, stärker bevorzugt mindestens etwa 70 bis 80 %, 80 bis 90 %, 90 bis 95 % und noch stärker bevorzugt mindestens etwa 96 %, 20 97 %, 98 %, 99 % oder noch homologer zu einer der Aminosäuresequenzen der SEQ ID NO: 2, 4, 6 oder 18 ist. Die erfindungsgemäße bevorzugte Desaturase besitzt vorzugsweise auch mindestens eine der hier beschriebenen Desaturase-Aktivitäten. Zum Beispiel umfasst eine erfindungsgemäße bevorzugte Desaturase eine Amino- 25 säuresequenz, die von einer Nukleotidsequenz kodiert wird, die, zum Beispiel unter stringenten Bedingungen, an eine Nukleotidsequenz der SEQ ID NO: 1, 3, 5 oder 11 hybridisiert und am Stoffwechsel von zum Aufbau von Zellmembranen in Phaeodactylum tricornutum notwendigen Verbindungen oder am Transport von Molekülen 30 35 über diese Membranen teilnehmen kann oder eine Doppelbindung in eine Fettsäure mit ein, zwei, drei oder vier Doppelbindungen und einer Kettenlänge von C₁₈, C₂₀ oder C₂₂ einführt.

Bei anderen Ausführungsformen ist die Desaturase im wesentlichen 40 homolog zu einer Aminosäuresequenz der SEQ ID NO: 2, 4 oder 6 und behält die funktionelle Aktivität des Proteins einer der Sequenzen der SEQ ID NO: 2, 4 oder 6 bei, ihre Aminosäuresequenz unterscheidet sich jedoch aufgrund von natürlicher Variation oder Mutagenese, wie eingehend im obigen Unterabschnitt I beschrieben. 45 Bei einer weiteren Ausführungsform ist die Desaturase folglich ein Protein, das eine Aminosäuresequenz umfasst, die mindestens etwa 50 bis 60 %, vorzugsweise mindestens etwa 60 bis 70 % und

62

stärker bevorzugt mindestens etwa 70 bis 80 %, 80 bis 90 %, 90 bis 95 % und am stärksten bevorzugt mindestens etwa 96 %, 97 %, 98 %, 99 % oder noch homologer zu einer vollständigen Aminosäuresequenz der SEQ ID NO: 2, 4 oder 6 ist und zumindest eine der hier beschriebenen Desaturase-Aktivitäten aufweist. Bei einer anderen Ausführungsform betrifft die Erfindung ein vollständiges Phaeodactylum tricornutum-Protein, das im wesentlichen homolog zu einer vollständigen Aminosäuresequenz der SEQ ID NO: 2, 4 oder 6 ist.

10

Biologisch aktive Teile einer Desaturase umfassen Peptide, umfassend Aminosäuresequenzen, die von der Aminosäuresequenz einer Desaturase hergeleitet sind, z.B. eine in SEQ ID NO: 2, 4 oder 6 gezeigte Aminosäuresequenz oder die Aminosäuresequenz eines Proteins, das zu einer Desaturase homolog ist, welche weniger Aminosäuren als die Vollängen-Desaturase oder das Vollängenprotein aufweisen, das zu einer Desaturase homolog ist, und zumindest eine Aktivität einer Desaturase aufweisen. Gewöhnlich umfassen biologisch aktive Teile (Peptide, z.B.

20 Peptide, die zum Beispiel 5, 10, 15, 20, 30, 35, 36, 37, 38, 39, 40, 50, 100 oder mehr Aminosäuren lang sind) eine Domäne oder ein Motiv mit mindestens einer Aktivität einer Desaturase. Überdies können andere biologisch aktive Teile, in denen andere Bereiche des Proteins deletiert sind, durch rekombinante 25 Techniken hergestellt und bezüglich einer oder mehrerer der hier beschriebenen Aktivitäten untersucht werden. Die biologisch aktiven Teile einer Desaturase umfassen vorzugsweise ein/eine oder mehrere ausgewählte Domänen/Motive oder Teile davon mit biologischer Aktivität.

30

Desaturasen werden vorzugsweise durch DNA-Rekombinationstechniken hergestellt. Zum Beispiel wird ein das Protein kodierendes Nukleinsäuremolekül in einen Expressionsvektor (wie vorstehend beschrieben) kloniert, der Expressionsvektor wird in eine 35 Wirtszelle (wie vorstehend beschrieben) eingebracht, und die Desaturase wird in der Wirtszelle exprimiert. Die Desaturase kann dann durch ein geeignetes Reinigungsschema mittels Standard-Proteinreinigungstechniken aus den Zellen isoliert werden. Alternativ zur rekombinannten Expression kann eine Desaturase, ein 40 -Polypeptid, oder -Peptid mittels Standard-Peptidsynthese-techniken chemisch synthetisiert werden. Überdies kann native Desaturase aus Zellen (z.B. Endothelzellen) z.B. unter Verwendung eines Anti-Desaturase-Antikörpers isoliert werden, der durch Standardtechniken produziert werden kann, wobei eine erfindungs-45 gemäße Desaturase oder ein Fragment davon verwendet wird.

Die Erfindung stellt auch chimäre Desaturase-Proteine oder Desaturase-Fusionsproteine bereit. Wie hier verwendet, umfasst ein "chimäres Desaturase-Protein" oder "Desaturase-Fusionsprotein" ein Desaturase-Polypeptid, das funktionsfähig an ein 5 nicht-Desaturase-Polypeptid gebunden ist. Ein "Desaturase-Polypeptid" betrifft ein Polypeptid mit einer Aminosäuresequenz, die einer Desaturase entspricht, wohingegen ein "nicht-Desaturase-Polypeptid" ein Polypeptid mit einer Aminosäuresequenz betrifft, die einem Protein entspricht, das im wesentlichen nicht homolog 10 zu der Desaturase ist, z.B. ein Protein, das sich vom der Desaturase unterscheidet und aus dem gleichen oder einem anderen Organismus stammt. Innerhalb des Fusionsproteins soll der Begriff "funktionsfähig verbunden" bedeuten, dass das Desaturase-Polypeptid und das nicht-Desaturase-Polypeptid 15 so miteinander fusioniert sind, dass beide Sequenzen die vorhergesagte, der verwendeten Sequenz zugeschriebene Funktion erfüllen. Das nicht-Desaturase-Polypeptid kann an den N-Terminus oder den C-Terminus des Desaturase-Polypeptids fusioniert sein. Bei einer Ausführungsform ist das Fusionsprotein zum Beispiel 20 ein GST-Desaturase-Fusionsprotein, bei dem die Desaturase-Sequenzen an den C-Terminus der GST-Sequenzen fusioniert sind. Diese Fusionsproteine können die Reinigung der rekombinannten Desaturasen erleichtern. Bei einer weiteren Ausführungsform ist das Fusionsprotein eine Desaturase, die eine heterologe Signal- 25 sequenz an ihrem N-Terminus aufweist. In bestimmten Wirtszellen (z.B. Säuger-Wirtszellen) kann die Expression und/oder Sekretion einer Desaturase durch Verwendung einer heterologen Signalsequenz gesteigert werden.

30 Ein erfindungsgemäßes chimäres Desaturase-Protein oder Desaturase-Fusionsprotein wird durch Standard-DNA-Rekombinations-techniken hergestellt. Zum Beispiel werden DNA-Fragmente, die unterschiedliche Polypeptidsequenzen kodieren, gemäß herkömmlicher Techniken im Leseraster aneinander ligiert, indem 35 beispielsweise glatte oder überhängende Enden zur Ligation, Restriktionsenzymspaltung zur Bereitstellung geeigneter Enden, Auffüllen kohäsiver Enden, wie erforderlich, Behandlung mit alkalischer Phosphatase, um ungewollte Verknüpfungen zu vermeiden, und enzymatische Ligation eingesetzt werden. Bei einer 40 weiteren Ausführungsform kann das Fusionsgen durch herkömmliche Techniken, einschließlich DNA-Syntheseautomaten, synthetisiert werden. Alternativ kann eine PCR-Amplifizierung von Genfragmenten unter Verwendung von Ankerprimern durchgeführt werden, die komplementäre Überhänge zwischen aufeinanderfolgenden Gen- 45 fragmenten erzeugen, die anschließend miteinander hybridisiert und reamplifiziert werden können, so dass eine chimäre Gensequenz erzeugt wird (siehe zum Beispiel Current Protocols in Molecular

Biology, Hrsgb. Ausubel et al., John Wiley & Sons: 1992). Überdies sind viele Expressionsvektoren kommerziell erhältlich, die bereits eine Fusionseinheit (z.B. ein GST-Polypeptid) kodieren. Eine Desaturase-kodierende Nukleinsäure kann in einen solchen 5 Expressionsvektor kloniert werden, so dass die Fusionseinheit im Leseraster mit dem Desaturase-Protein verbunden ist.

Homologe der Desaturase können durch Mutagenese, z.B. durch spezifische Punktmutation oder Verkürzung der Desaturase, erzeugt 10 werden. Der Begriff "Homologe", wie hier verwendet, betrifft eine variante Form der Desaturase, die als Agonist oder Antagonist der Desaturase-Aktivität wirkt. Ein Agonist der Desaturase kann im wesentlichen die gleiche Aktivität wie die oder einen Teil der biologischen Aktivitäten der Desaturase beibehalten. Ein 15 Antagonist der Desaturase kann eine oder mehrere Aktivitäten der natürlich vorkommenden Form der Desaturase durch zum Beispiel kompetitive Bindung an ein stromabwärts oder -aufwärts gelegenes Element der Stoffwechselkaskade für Zellmembrankomponenten, welche die Desaturase umfasst, oder durch Bindung an eine 20 Desaturase, welche den Transport von Verbindungen über Zellmembranen vermittelt, hemmen, wodurch die Translokation gehemmt wird.

Bei einer alternativen Ausführungsform können Homologe der Desaturase durch Sichten kombinatorischer Banken von Mutanten, 25 z.B. Verkürzungsmutanten, der Desaturase hinsichtlich Desaturase-Agonisten- oder -Antagonisten-Aktivität identifiziert werden. Bei einer Ausführungsform wird eine variegierte Bank von Desaturase-Varianten durch kombinatorische Mutagenese auf Nukleinsäure-ebene erzeugt und durch eine variegierte Genbank kodiert. Eine 30 variegierte Bank von Desaturase-Varianten kann z.B. durch enzymatische Ligation eines Gemisches von synthetischen Oligonukleotiden in Gensequenzen hergestellt werden, so dass sich ein degenerierter Satz potentieller Desaturase-Sequenzen als individuelle Polypeptide oder alternativ als Satz größerer 35 Fusionsproteine (z.B. für das Phage-Display), die diesen Satz von Desaturase-Sequenzen enthalten, exprimieren lässt. Es gibt eine Vielzahl von Verfahren, die zur Herstellung von Banken potentieller Desaturase-Homologen aus einer degenerierten Oligonukleotidsequenz verwendet werden können. Die chemische 40 Synthese einer degenerierten Gensequenz kann in einem DNA-Synthesearmaten durchgeführt und das synthetische Gen dann in einen geeigneten Expressionsvektor ligiert werden. Die Verwendung eines degenerierten Satzes von Genen ermöglicht die Bereitstellung sämtlicher Sequenzen, die den gewünschten Satz 45 an potentiellen Desaturase-Sequenzen kodieren, in einem Gemisch. Verfahren zur Synthese degenerierter Oligonukleotide sind im Fachgebiet bekannt (siehe z.B. Narang, S.A. (1983) Tetrahedron

39:3; Itakura et al. (1984) Annu. Rev. Biochem. 53:323; Itakura et al., (1984) Science 198:1056; Ike et al. (1983) Nucleic Acids Res. 11:477).

5 Zusätzlich können Banken von Desaturase-Fragmenten zur Herstellung einer variierten Population von Desaturase-Fragmenten für das Sichten und für die anschließende Selektion von Homologen einer Desaturase verwendet werden. Bei einer Ausführungsform kann eine Bank von Fragmenten der kodierenden Sequenz durch
10 Behandeln eines doppelsträngigen PCR-Fragmentes einer kodierenden Desaturase-Sequenz mit einer Nuklease unter Bedingungen, unter denen Doppelstrangbrüche nur etwa einmal pro Molekül erfolgen, Denaturieren der doppelsträngigen DNA, Renaturieren der DNA unter Bildung doppelsträngiger DNA, welche Sense/Antisense-Paare von
15 verschiedenen Produkten mit Doppelstrangbrüchen umfassen kann, Entfernen einzelsträngiger Abschnitte aus neu gebildeten Duplices durch Behandlung mit S1-Nuklease und Ligieren der resultierenden Fragmentbank in einen Expressionsvektor erzeugt werden. Mit diesem Verfahren kann eine Expressionsbank hergeleitet werden,
20 die N-terminale, C-terminale und interne Fragmente der Desaturase verschiedener Größen kodiert.

Im Fachgebiet sind mehrere Techniken für das Sichten von Genprodukten in kombinatorischen Banken, die durch Punktmutationen oder Verkürzung hergestellt worden sind, und für das Sichten von cDNA-Banken nach Genprodukten mit einer ausgewählten Eigenschaft bekannt. Diese Techniken lassen sich an das schnelle Sichtung der Genbanken anpassen, die durch kombinatorische Mutagenese von Desaturase-Homologen erzeugt worden sind. Die am häufigsten verwendeten Techniken zum Sichtung großer Genbanken, die einer Analyse mit hohem Durchsatz unterworfen werden können, umfassen gewöhnlich das Klonieren der Genbank in replizierbare Expressionsvektoren, Transformieren von geeigneten Zellen mit der resultierenden Vektorenbank und Exprimieren der kombinatorischen Gene unter Bedingungen, unter denen der Nachweis der gewünschten Aktivität die Isolation des Vektors, der das Gen kodiert, dessen Produkt nachgewiesen wurde, erleichtert. Recursive-Ensemble-Mutagenese (REM), eine neue Technik, die die Häufigkeit funktioneller Mutanten in den Banken erhöht, kann in Kombination mit den Sichtungstests zur Identifikation von Desaturase-Homologen verwendet werden (Arkin und Yourvan (1992) Proc. Natl. Acad. Sci. USA 89:7811-7815; Delgrave et al. (1993) Protein Engineering 6(3):327-331).
45 Eine weitere bekannte Technik zur Veränderung von katalytischen Eigenschaften von Enzymen bzw. deren codierenden Genen ist das "Gen-Shuffling" (siehe z.B. in Stemmer, PNAS 1994, 91:

10747-10751, WO9720078 oder WO9813487), das eine Kombination von Genfragmenten darstellt, wobei diese Neukombination zusätzlich noch durch fehlerhafte Polymerasenkettenreaktionen variiert werden kann und somit eine hohe zu testende Sequenzdiversität schafft.

5 Voraussetzung für den Einsatz eines solchen Ansatzes ist jedoch ein geeignetes Screeningsystem, um die erstellte Gendiversität auf Funktionalität zu überprüfen.

Insbesondere für die Sichtung von Desaturaseaktivitäten

10 ist ein Sichtungsverfahren Voraussetzung, das PUFA-abhängig Enzymaktivität(en) erfaßt. Bzgl. Desaturaseaktivitäten mit Spezifität für PUFAs kann man in Mucor-Species, die durch bekannte Transformationsverfahren mit gewünschten Genkonstrukten transformierbar sind, die Toxizität von Arachidonsäure in An-

15 wesenheit eines toxischen Metaboliten (hier: Salicylsäure oder Salicylsäurederivate) nutzen (Eroshin et al., Mikrobiologiya, Vol. 65, No.1 1996, Seiten 31-36), um eine wachstumsbasierte Erstsichtung durchzuführen. Resultierende Klone können dann einer Analyse ihrer Lipidinhaltstoffe mittels Gaschromatographie und

20 Massenspektroskopie unterzogen werden, um Edukte und Produkte in Art und Menge zu erfassen.

Bei einer weiteren Ausführungsform können Tests auf Zellbasis zur Analyse einer variegierten Desaturase-Bank unter Verwendung

25 von weiteren im Fachgebiet bekannten Verfahren ausgenutzt werden.

E. Erfindungsgemäße Verwendungen und Verfahren

Die hier beschriebenen Nukleinsäuremoleküle, Proteine, Protein-

30 homologen, Fusionsproteine, Primer, Vektoren und Wirtszellen können bei einem oder mehreren der nachstehenden Verfahren verwendet werden: Identifikation von Phaeodactylum und verwandten Organismen, Kartierung der Genome von Organismen, die mit Phaeodactylum tricornutum verwandt sind, Identifikation und

35 Lokalisierung von Phaeodactylum tricornutum-Sequenzen von Interesse, Evolutionsstudien, Bestimmung von Desaturase-Proteinbereichen, die für die Funktion notwendig sind, Modulation einer Desaturase-Aktivität; Modulation des Stoffwechsels einer oder mehrerer Zellmembrankomponenten; Modulation des Transmembran-

40 transports einer oder mehrerer Verbindungen sowie Modulation der zellulären Produktion einer gewünschten Verbindung, wie einer Feinchemikalie. Die erfundungsgemäßen Desaturase-Nukleinsäuremoleküle haben eine Vielzahl von Verwendungen. Sie können zunächst zur Identifikation eines Organismus als Phaeodactylum

45 tricornutum oder als naher Verwandter davon verwendet werden. Sie können auch zur Identifikation des Vorliegens von Phaeodactylum tricornutum oder eines Verwandten davon in einer Mischpopulation

von Mikroorganismen verwendet werden. Die Erfindung stellt die Nukleinsäuresequenzen einer Reihe von *Phaeodactylum tricornutum*-Genen bereit; durch Sondieren der extrahierten genomischen DNA einer Kultur einer einheitlichen oder gemischten Population von 5 Mikroorganismen unter stringenten Bedingungen mit einer Sonde, die einen Bereich eines *Phaeodactylum tricornutum*-Gens oder von Teilen davon überspannt, das für diesen Organismus einzigartig ist, kann man bestimmen, ob dieser Organismus vorliegt. *Phaeodactylum tricornutum* selbst werden zur kommerziellen Produktion 10 mehrfach ungesättigter Säuren verwendet und eignen darüber hinaus zur PUFA-Produktion auch in anderen Organismen insbesondere wenn erreicht werden soll, dass resultierende PUFAs auch in die Triacylglycerolfraktion eingebaut werden sollen.

15 Ferner können die erfindungsgemäßen Nukleinsäure- und Proteinkomplexe als Marker für spezifische Bereiche des Genoms dienen. Dies ist nicht nur zur Kartierung des Genoms, sondern auch für funktionelle *Phaeodactylum tricornutum*-Proteinen geeignet. Zur Identifikation des Genombereichs, an den ein bestimmtes DNA- 20 bindendes Protein von *Phaeodactylum tricornutum* bindet, könnte das *Phaeodactylum tricornutum*-Genom zum Beispiel gespalten werden und die Fragmente mit dem DNA-bindenden Protein inkubiert werden. Diejenigen, die das Protein binden, können zusätzlich mit den erfindungsgemäßen Nukleinsäuremolekülen, vorzugsweise mit leicht 25 nachweisbaren Markierungen, sondiert werden; die Bindung eines solchen Nukleinsäuremoleküls an das Genomfragment ermöglicht die Lokalisierung des Fragments auf der Genomkarte von *Phaeodactylum tricornutum* und erleichtert, wenn dies mehrmals mit unterschiedlichen Enzymen durchgeführt wird, eine rasche Bestimmung der 30 Nukleinsäuresequenz, an die das Protein bindet. Die erfindungsgemäßen Nukleinsäuremoleküle können zudem ausreichend homolog zu den Sequenzen verwandter Arten sein, dass diese Nukleinsäuremoleküle als Marker für die Konstruktion einer genomischen Karte bei verwandten Pilzen oder Algen dienen können.

35 Die erfindungsgemäßen Desaturase-Nukleinsäuremoleküle eignen sich auch für Evolutions- und Proteinstruktur-Untersuchungen. Die Stoffwechsel- und Transportprozesse, an denen die erfindungsgemäßen Moleküle beteiligt sind, werden von vielen 40 prokaryotischen und eukaryotischen Zellen genutzt; durch Vergleich der Sequenzen der erfindungsgemäßen Nukleinsäuremoleküle mit solchen, die ähnliche Enzyme aus anderen Organismen kodieren, kann der Evolutions-Verwandschaftsgrad der Organismen bestimmt werden. Entsprechend ermöglicht ein solcher Vergleich 45 die Bestimmung, welche Sequenzbereiche konserviert sind und welche nicht, was bei der Bestimmung von Bereichen des Proteins hilfreich sein kann, die für die Enzymfunktion essentiell sind.

Dieser Typ der Bestimmung ist für Proteinengineering-Untersuchungen wertvoll und kann einen Hinweis darauf geben, wieviel Mutagenese das Protein tolerieren kann, ohne die Funktion zu verlieren.

5

Die Manipulation der erfindungsgemäßen Desaturase-Nukleinsäuremoleküle kann zur Produktion von Desaturasen mit funktionellen Unterschieden zu den Wildtyp-Desaturasen führen. Die Effizienz oder Aktivität dieser Proteine kann verbessert sein, sie können 10 in größeren Anzahlen als gewöhnlich in der Zelle zugegen sein, oder ihre Effizienz oder Aktivität kann verringert sein. Verbesserte Effizienz oder Aktivität bedeutet beispielsweise, dass das Enzym eine höhere Selektivität und/oder Aktivität, vorzugsweise eine mindestens 10 % höhere, besonders bevorzugt eine 15 mindestens 20 % höhere Aktivität, ganz besonders bevorzugt eine mindestens 30 % höhere Aktivität als das ursprüngliche Enzym aufweist.

Es gibt eine Reihe von Mechanismen, durch die die Veränderung 20 einer erfindungsgemäßen Desaturase die Ausbeute, Produktion und/oder Effizienz der Produktion einer Feinchemikalie, welche ein solches verändertes Protein enthält, direkt beeinflussen kann. Die Gewinnung von Feinchemikalien-Verbindungen aus Kulturen von Ciliaten, Algen oder Pilzen im großen Maßstab ist signifikant verbessert, wenn die Zelle die gewünschten Verbindungen sezerniert, da diese Verbindungen aus dem Kulturmedium (im Gegensatz zur Extraktion aus der Masse der gezüchteten Zellen) leicht gereinigt werden können. Ansonsten lässt sich die Reinigung verbessern, wenn die Zelle in vivo Verbindungen in einem spezialisierten Kompartiment mit einer Art Konzentrationsmechanismus speichert. Bei Pflanzen, die Desaturasen exprimieren, kann ein gesteigerter Transport zu besserer Verteilung innerhalb des Pflanzengewebes und der -organe führen. Durch Vergrößern der Anzahl oder der Aktivität von Transportermolekülen, welche Feinchemikalien aus der Zelle exportieren, kann es möglich sein, die 35 Menge der produzierten Feinchemikalie, die im extrazellulären Medium zugegen ist, zu steigern, wodurch Ernte und Reinigung oder bei Pflanzen eine effizientere Verteilung erleichtert werden. Zur effizienten Überproduktion einer oder mehrerer Feinchemikalien 40 sind dagegen erhöhte Mengen an Cofaktoren, Vorläufermolekülen und Zwischenverbindungen für die geeigneten Biosynthesewege erforderlich. Durch Vergrößern der Anzahl und/oder der Aktivität von Transporterproteinen, die am Import von Nährstoffen, wie Kohlenstoffquellen (d.h. Zuckern), Stickstoffquellen (d.h. Aminosäuren, 45 Ammoniumsalzen), Phosphat und Schwefel, beteiligt sind, kann man die Produktion einer Feinchemikalie aufgrund der Beseitigung aller Einschränkungen des Nährstoffangebots beim Biosynthese-

prozess verbessern. Fettsäuren, wie PUFAs, und Lipide, die PUFAs enthalten, sind selbst wünschenswerte Feinchemikalien; durch Optimieren der Aktivität oder Erhöhen der Anzahl einer oder mehrerer erfindungsgemäßer Desaturasen, die an der Biosynthese dieser Verbindungen beteiligt sind, oder durch Stören der Aktivität einer oder mehrerer Desaturasen, die am Abbau dieser Verbindungen beteiligt sind, kann man somit die Ausbeute, Produktion und/oder Effizienz der Produktion von Fettsäure- und Lipidmoleküle in Ciliaten, Algen, Pflanzen, Pilzen, Hefen oder anderen Mikroorganismen steigern.

Die Manipulation eines oder mehrerer erfindungsgemäßer Desaturase-Gene kann ebenfalls zu Desaturasen mit veränderten Aktivitäten führen, welche die Produktion einer oder mehrerer gewünschter Feinchemikalien aus Algen, Pflanzen, Ciliaten oder Pilzen indirekt beeinflussen. Die normalen biochemischen Stoffwechselprozesse führen z.B. zur Produktion einer Vielzahl an Abfallprodukten (z.B. Wasserstoffperoxid und andere reaktive Sauerstoffspezies), die diese Stoffwechselprozesse aktiv stören können (z.B. nitriert Peroxynitrit bekanntlich Tyrosin-Seitenketten, wodurch einige Enzyme mit Tyrosin im aktiven Zentrum inaktiviert werden (Groves, J.T. (1999) Curr. Opin. Chem. Biol. 3 (2);226-235)). Diese Abfallprodukte werden zwar üblicherweise ausgeschieden, aber die zur fermentativen Produktion im großen Maßstab verwendeten Zellen werden für die Überproduktion einer oder mehrerer Feinchemikalien optimiert und können somit mehr Abfallprodukte produzieren als für eine Wildtypzelle üblich. Durch Optimieren der Aktivität einer oder mehrerer erfindungsgemäßer Desaturasen, die am Export von Abfallmolekülen beteiligt sind, kann man die Lebensfähigkeit der Zelle verbessern und eine effiziente Stoffwechselaktivität aufrechterhalten. Auch das Vorliegen hoher intrazellulärer Mengen der gewünschten Feinchemikalie kann tatsächlich für die Zelle toxisch sein, so dass man durch Steigern der Fähigkeit der Zelle zur Sekretion dieser Verbindungen die Lebensfähigkeit der Zelle verbessern kann.

Die erfindungsgemäßen Desaturasen können ferner so manipuliert sein, dass die relativen Mengen verschiedener Lipid- und Fett-säuremoleküle verändert werden. Dies kann eine entscheidende Auswirkung auf die Lipidzusammensetzung der Zellmembran haben. Da jeder Lipidtyp unterschiedliche physikalische Eigenschaften hat, kann eine Veränderung der Lipidzusammensetzung einer Membran die Membranfluidität signifikant verändern. Änderungen der Membranfluidität können den Transport von Molekülen über die Membran beeinflussen, was, wie vorstehend erläutert, den Export von Abfallprodukten oder der produzierten Feinchemikalie oder den Import notwendiger Nährstoffe modifizieren kann. Diese Änderungen

der Membranfluidität können auch die Integrität der Zelle entscheidend beeinflussen; Zellen mit vergleichsweise schwächeren Membranen sind anfälliger gegenüber abiotischen und biotischen Stressbedingungen, welche die Zelle beschädigen oder abtöten
5 können. Durch Manipulieren von Desaturasen, die an der Produktion von Fettsäuren und Lipiden für den Membranaufbau beteiligt sind, so dass die resultierende Membran eine Membranzusammensetzung hat, die für die in den Kulturen, die zur Produktion von Feinchemikalien verwendet werden, herrschenden Umweltbedingungen
10 empfänglicher sind, sollte ein größerer Anteil der Zellen überleben und sich vermehren. Größere Mengen an produzierenden Zellen sollten sich in größeren Ausbeuten, höherer Produktion oder Effizienz der Produktion der Feinchemikalie aus der Kultur manifestieren.
15

Die vorstehend genannten Mutagenesestrategien für Desaturasen, die zu erhöhten Ausbeuten einer Feinchemikalie führen sollen, sollen nicht beschränkend sein; Variationen dieser Strategien sind dem Fachmann leicht ersichtlich. Unter Verwendung dieser
20 Mechanismen und mithilfe der hier offenbarten Mechanismen können die erfindungsgemäßen Nukleinsäure- und Proteinmoleküle zur Erzeugung von Algen, Ciliaten, Pflanzen, Tieren, Pilzen oder anderen Mikroorganismen, wie *C. glutamicum*, verwendet werden, die mutierte Desaturase-Nukleinsäure- und Proteinmoleküle exprimieren, so dass die Ausbeute, Produktion und/oder Effizienz der Produktion einer gewünschten Verbindung verbessert wird. Diese gewünschte Verbindung kann ein beliebiges natürliches Produkt von Algen, Ciliaten, Pflanzen, Tieren, Pilzen oder Bakterien sein, welches die Endprodukte von Biosynthesewegen und Zwischenprodukte
25 30 natürlich vorkommender Stoffwechselwege umfasst, sowie Moleküle, die im Stoffwechsel dieser Zellen nicht natürlich vorkommen, die jedoch von den erfindungsgemäßen Zellen produziert werden.

Eine weitere erfindungsgemäße Ausführungsform ist ein Verfahren
35 zur Produktion von PUFAs, wobei das Verfahren das Züchten eines Organismus, der eine erfindungsgemäße Nukleinsäure, ein erfindungsgemäßes Genkonstrukt oder einen erfindungsgemäßen Vektor umfasst, welche ein Polypeptid kodieren, das C₁₈-, C₂₀- oder C₂₂-Fettsäuren mit mindestens zwei Doppelbindungen im
40 Fettsäuremolekül um mindestens zwei Kohlenstoffatome unter Bedingungen, unter denen PUFAs in dem Organismus produziert werden, verlängert, umfasst. Durch dieses Verfahren hergestellte PUFAs lassen sich durch Ernten der Organismen entweder aus der Kultur, in der sie wachsen, oder von dem Feld, Aufbrechen
45 und/oder Extrahieren des geernteten Materials mit einem organischen Lösungsmittel isolieren. Aus diesem Lösungsmittel kann das Öl, das Lipide, Phospholipide, Sphingolipide, Glyco-

lipide, Triacylglycerine und/oder freie Fettsäuren mit höherem Gehalt an PUFAs enthält, isoliert werden. Durch basische oder saure Hydrolyse der Lipide, Phospholipide, Sphingolipide, Glycolipide, Triacylglycerine können die freien Fettsäuren mit höherem Gehalt an PUFAs isoliert werden. Ein höherer Gehalt an PUFAs bedeutet mindestens 5 %, vorzugsweise 10 %, besonders bevorzugt 20 %, ganz besonders bevorzugt 40 % mehr PUFAs als der ursprüngliche Organismus, der keine zusätzliche Nukleinsäure, die die erfundungsgemäße Desaturase kodiert, besitzt.

10

Vorzugsweise sind die durch dieses Verfahren produzierten PUFAs C₁₈- oder C₂₀₋₂₂-Fettsäuremoleküle mit mindestens zwei Doppelbindungen im Fettsäuremolekül, vorzugsweise drei, vier, bei Kombination mit einer weiteren Elongasen und einer Δ-4 Desaturase fünf oder sechs Doppelbindungen. Diese C₁₈- oder C₂₀₋₂₂-Fettsäuremoleküle lassen sich aus dem Organismus in Form eines Öls, Lipids oder einer freien Fettsäure isolieren. Geeignete Organismen sind beispielsweise die vorstehend erwähnten. Bevorzugte Organismen sind transgene Pflanzen.

20

Eine erfundungsgemäße Ausführungsform sind Öle, Lipide oder Fettsäuren oder Fraktionen davon, die durch das oben beschriebene Verfahren hergestellt worden sind, besonders bevorzugt Öl, Lipid oder eine Fettsäurezusammensetzung, die PUFAs umfassen und von transgenen Pflanzen herrühren.

Eine weitere erfundungsgemäße Ausführungsform ist die Verwendung des Öls, Lipids oder der Fettsäurezusammensetzung in Futtermitteln, Nahrungsmitteln, Kosmetika oder Pharmazeutika.

30

Ein weiterer Erfindungsgegenstand ist ein Verfahren zur Identifikation eines Antagonisten oder Agonisten von Desaturasen, umfassend

35 a)

in Kontaktbringen der Zellen, die das Polypeptid der vorliegenden Erfindung exprimieren, mit einem Kandidatenstoff;

b) Testen der Desaturaseaktivität;

40

c) Vergleichen der Desaturaseaktivität mit einer Standardaktivität in Abwesenheit des Kandidatenstoffs, wobei ein Anstieg der Desaturaseaktivität über den Standard anzeigt, daß der Kandidatenstoff ein Agonist und eine Verringerung der Desaturaseaktivität anzeigt, daß der Kandidatenstoff ein Antagonist ist.

Der genannte Kandidatenstoff kann ein chemisch synthetisierter oder mikrobiologisch produzierter Stoff sein und z.B. in Zellextrakten von z.B. Pflanzen, Tieren oder Mikroorganismen auftreten. Weiterhin kann der genannte Stoff zwar im Stand der Technik bekannt sein, aber bisher nicht bekannt sein als die Aktivität der Desaturasen steigernd oder reprimierend. Das Reaktionsgemisch kann ein zellfreier Extrakt sein oder eine Zelle oder Zellkultur umfassen. Geeignete Methoden sind dem Fachmann bekannt und werden z.B. allgemein beschrieben in Alberts, Molecular Biology the cell, 3rd Edition (1994), z.B. Kapitel 17. Die genannten Stoffe können z.B. zu dem Reaktionsgemisch oder dem Kulturmedium zugegeben werden oder den Zellen injiziert werden oder auf eine Pflanze gesprüht werden.

Wenn eine Probe, die ein nach der erfindungsgemäßen Methode aktiven Stoff beinhaltet, identifiziert wurde, dann ist es entweder möglich, den Stoff direkt von der ursprünglichen Probe zu isolieren oder man kann die Probe in verschiedene Gruppen teilen, z.B. wenn sie aus einer Vielzahl von verschiedenen Komponenten besteht, um so die Zahl der verschiedenen Substanzen pro Probe zu reduzieren und dann das erfindungsgemäße Verfahren mit einer solchen "Unterprobe" der ursprünglichen Probe zu wiederholen. Abhängig von der Komplexität der Probe können die oben beschriebenen Schritte mehrmals wiederholt werden, vorzugsweise bis die gemäß der erfindungsgemäßen Methode identifizierte Probe nur noch eine geringe Anzahl von Substanzen oder nur noch eine Substanz umfaßt. Vorzugsweise wird der gemäß der erfindungsgemäßen Methode identifizierte Stoff oder Derivate davon weiter formuliert, so, daß er für die Anwendung in der Pflanzenzüchtung oder Pflanzenzell- oder Gewebekultur geeignet ist.

Die Stoffe, die gemäß dem erfindungsgemäßen Verfahren getestet und identifiziert wurden, können sein: Expressionsbibliotheken, z.B. cDNA-Expressionsbibliotheken, Peptide, Proteine, Nukleinsäuren, Antikörper, kleine organische Stoffe, Hormone, PNAs oder ähnliches (Milner, Nature Medicin 1 (1995), 879-880; Hupp, Cell. 83 (1995), 237-245; Gibbs, Cell. 79 (1994), 193-198 und darin zitierte Referenzen). Diese Stoffe könne auch funktionelle Derivate oder Analogon der bekannten Inhibitoren oder Aktivatoren sein. Verfahren zur Herstellung von chemischen Derivaten oder Analogon sind dem Fachmann bekannt. Die genannten Derivate und Analogon können gemäß Verfahren nach dem Stand der Technik getestet werden. Weiterhin kann computergestütztes Design oder Peptidomimetics zur Herstellung geeigneter Derivate und Analogon verwendet werden. Die Zelle oder das Gewebe, die/das für das erfindungsgemäße Verfahren verwendet werden kann, ist vorzugsweise eine erfindungsgemäße Wirtszelle, Pflanzenzelle oder ein

Pflanzengewebe, wie in den oben genannten Ausführungsformen beschrieben.

Entsprechend betrifft die vorliegende Erfindung auch einen
5 Stoff, der gemäß den vorstehenden erfindungsgemäßen Verfahren identifiziert wurde. Der Stoff ist z.B. ein Homolog der erfindungsgemäßen Desaturasen. Homologe der Desaturasen können durch Mutagenese, z.B. durch Punktmutation oder Deletion der Desaturasen, erzeugt werden. Hierin verwendet wird der Begriff
10 "Homolog" als eine variante Form der Desaturasen, die als Agonist oder Antagonist für die Aktivität der Desaturasen wirkt. Ein Agonist kann die im wesentlichen gleiche oder einen Teil der biologischen Aktivität der Desaturasen haben. Ein Antagonist der Desaturasen kann eine oder mehr Aktivitäten der natürlich vor-
15 kommenden Formen der Desaturasen inhibieren, z.B. kompetitiv an ein Downstream oder Upstream gelegenes Mitglied der Fettsäure-synthese-Stoffwechselwege, die die Desaturasen einschließen, binden oder an Desaturasen binden und dabei die Aktivität reduzieren oder inhibieren.

20 Außerdem betrifft die vorliegende Erfindung auch ein Antikörper oder ein Fragment davon, wie sie hierin beschrieben werden, der die Aktivität der erfindungsgemäßen Desaturasen inhibiert.

25 Bei einem Aspekt betrifft die vorliegende Erfindung ein Antikörper, der spezifisch den erfindungsgemäßen oben beschriebenen Agonisten oder Antagonisten erkennt bzw. bindet.

Ein weiterer Aspekt betrifft eine Zusammensetzung, die den Antikörper, den nach dem erfindungsgemäßen Verfahren identifizierten Stopp oder das Antisense-Molekül umfaßt.

In einer weiteren Ausführungsform betrifft die vorliegende Erfindung ein Kit, umfassend die erfindungsgemäße Nukleinsäure, 35 das erfindungsgemäße Genkonstrukt, die erfindungsgemäße Aminosäuresequenz, das erfindungsgemäße Antisense-Nukleinsäuremolekül, den erfindungsgemäßen Antikörper und/oder Zusammensetzung, einen Antagonisten oder Agonisten, der nach dem erfindungsgemäßen Verfahren hergestellt wurde, und/oder erfindungsgemäße Öle, Lipide 40 und/oder Fettsäuren oder eine Fraktion davon. Ebenso kann das Kit die erfindungsgemäßen Wirtszellen, Organismen, Pflanzen oder Teile davon, erntbare Teile der erfindungsgemäßen Pflanzen oder Vermehrungsmaterial oder aber auch den erfindungsgemäßen Antagonisten oder Agonisten umfassen. Die Komponenten des Kits 45 der vorliegenden Erfindung können in geeigneten Containern, beispielsweise mit oder in Puffern oder anderen Lösungen verpackt sein. Ein oder mehr der genannten Komponenten können in ein und

demselben Container verpackt sein. Zusätzlich oder alternativ können ein oder mehr der genannten Komponenten z.B. auf einer festen Oberfläche adsorbiert sein, z.B. Nitrozellulosefilter, Glasplatten, Chips, Nylonmembranen oder Mikrotiterplatten. Das 5 Kit kann für jede der hierin beschriebenen Methoden und Ausführungsformen verwendet werden, z.B. für die Produktion von Wirtszellen, transgenen Pflanzen, zur Detektion von homologen Sequenzen, zur Identifikation von Antagonisten oder Agonisten usw. Weiterhin kann das Kit Anleitungen für die Verwendung des 10 Kits für eine der genannten Anwendungen enthalten.

Diese Erfindung wird durch die nachstehenden Beispiele weiter veranschaulicht, die nicht als beschränkend aufgefaßt werden sollten. Der Inhalt sämtlicher in dieser Patentanmeldung 15 zitierten Literaturstellen, Patentanmeldungen, Patente und veröffentlichten Patentanmeldungen ist hier durch Bezugnahme aufgenommen.

Beispielteil

20

Beispiel 1: Allgemeine Verfahren

a) Allgemeine Klonierungsverfahren:

25 Klonierungsverfahren, wie beispielsweise Restriktionsspaltungen, Agarosegelektrophorese, Reinigung von DNA-Fragmenten, Transfer von Nukleinsäuren auf Nitrocellulose- und Nylonmembranen, Verbindung von DNA-Fragmenten, Transformation von Escherichia coli- und Hefe-Zellen, Anzucht von Bakterien und Sequenzanalyse 30 rekombinanter DNA, wurden durchgeführt wie beschrieben in Sambrook et al. (1989) (Cold Spring Harbor Laboratory Press: ISBN 0-87969-309-6) oder Kaiser, Michaelis und Mitchell (1994) "Methods in Yeast Genetics" (Cold Spring Harbor Laboratory Press: ISBN 0-87969-451-3). Die Transformation und Anzucht von Algen, 35 wie Chlorella oder Phaeodactylum werden durchgeführt wie beschrieben von El-Sheekh (1999), Biologia Plantarum 42:209-216; Apt et al. (1996) Molecular and General Genetics 252 (5):872-9.

b) Chemikalien

40

Die verwendeten Chemikalien wurden, wenn im Text nicht anders angegeben, in p. A.-Qualität von den Firmen Fluka (Neu-Ulm), Merck (Darmstadt), Roth (Karlsruhe), Serva (Heidelberg) und Sigma (Deisenhofen) bezogen. Lösungen wurden unter Verwendung 45 von reinem pyrogenfreiem Wasser, im nachstehenden Text als H₂O bezeichnet, aus einer Milli-Q-Wassersystem-Wasserreinigungsanlage (Millipore, Eschborn) hergestellt. Restriktionsendo-

nukleasen, DNA-modifizierende Enzyme und molekularbiologische Kits wurden bezogen von den Firmen AGS (Heidelberg), Amersham (Braunschweig), Biometra (Göttingen), Boehringer (Mannheim), Genomed (Bad Oeynhausen), New England Biolabs (Schwalbach/Taunus), Novagen (Madison, Wisconsin, USA), Perkin-Elmer (Weiterstadt), Pharmacia (Freiburg), Qiagen (Hilden) und Stratagene (Amsterdam, Niederlande). Wenn nicht anders angegeben, wurden sie nach den Anweisungen des Herstellers verwendet.

10

c) Zellmaterial

Die erfindungsgemäßen isolierten Nukleinsäuresequenzen sind im Genom eines *Phaeodactylum tricornutum* UTEX646-Stammes enthalten, der über die Algensammlung der University of Texas, Austin verfügbar ist.

Phaeodactylum tricornutum wurde bei 25°C mit einem Licht/Dunkel Rhythmus von 14:10 Stunden bei 22°C und 35 microEinstein (entspricht micromol Photonen pro Quadratmeter und Sekunde) in Glasröhren kultiviert, die von unten mit Luft begast wurden.

Als Kulturmedium für *Phaeodactylum tricornutum* wurde das f/2 Kulturmedium mit 10 % organischen Medium nach Guillard, R.R.L. verwendet (1975; Culture of phytoplankton for feeding marine invertebrates. In: Smith, W.L. and Chanley, M.H. (Eds.) Culture of marine Invertebrate animals, NY Plenum Press, pp. 29-60.): Es enthält

30 995,5 ml Seewasser (artifiziell)

1 ml NaNO₃ (75 g/l), 1 ml NaH₂PO₄ (5 g/l), 1 ml Spurenelementelösung, 1 ml Tris/Cl pH 8.0, 0.5 ml f/2 Vitaminlösung

Spurenelementelösung: Na₂EDTA (4,36 g/l), FeCl₃ (3,15 g/l), Primäre Spurenelemente: CuSO₄ (10 g/l), ZnSO₄ (22 g/l), CoCl₂ (10 g/l), MnCl₂ (18 g/l), NaMoO₄ (6,3 g/l)

f/2 Vitaminlösung: Biotin: 10 mg/l, Thiamin 200 mg/l, Vit B12 0,1 mg/l

org-Medium: Na-Acetat (1 g/l), Glucose (6 g/l), Na-Succinat (3 g/l), Bacto-Trypton (4 g/l), Hefe-Extrakt (2 g/l)

Beispiel: 2 Isolierung von Gesamt-DNA aus *Phaeodactylum tricornutum* UTEX646 für Hybridisierungsexperimente

Die Einzelheiten der Isolierung von Gesamt-DNA betreffen die Aufarbeitung von Pflanzenmaterial mit einem Frischgewicht von einem Gramm.

CTAB-Puffer: 2 % (Gew./Vol.) N-Acetyl-N,N,N-trimethylammoniumbromid (CTAB); 100 mM Tris-HCl, pH 8,0; 1,4 M NaCl; 20 mM EDTA.

10

N-Laurylsarkosin-Puffer: 10 % (Gew./Vol.) N-Laurylsarkosin; 100 mM Tris-HCl, pH 8,0; 20 mM EDTA.

15 *Phaeodactylum tricornutum*-Zellmaterial wurde unter flüssigem Stickstoff in einem Mörser verrieben, so dass ein feines Pulver erhalten wurde, und in 2 ml-Eppendorfgefäß überführt. Das gefrorene Pflanzenmaterial wurde dann mit einer Schicht von 1 ml Zersetzungspuffer (1 ml CTAB-Puffer, 100 ml N-Laurylsarkosin-Puffer, 20 ml β -Mercaptoethanol und 10 ml Proteinase K-Lösung, 20 10 mg/ml) überschichtet und eine Stunde unter kontinuierlichem Schütteln bei 60°C inkubiert. Das erhaltene Homogenat wurde in zwei Eppendorfgefäß (2 ml) aufgeteilt und zweimal durch Schütteln mit dem gleichen Volumen Chloroform/Isoamylalkohol (24:1) extrahiert. Zur Phasentrennung wurde eine Zentrifugation 25 bei 8000 x g und RT (= Raumtemperatur = ~ 23°C) jeweils 15 min lang durchgeführt. Die DNA wurde dann 30 min unter Verwendung von eiskaltem Isopropanol bei -70°C gefällt. Die gefällte DNA wurde bei 10000 g 30 min bei 4°C sedimentiert und in 180 ml TE-Puffer (Sambrook et al., 1989, Cold Spring Harbor Laboratory Press: ISBN 0-87969-309-6) resuspendiert. Zur weiteren Reinigung wurde die DNA mit NaCl (1,2 M Endkonzentration) behandelt und erneut 30 min unter Verwendung des zweifachen Volumens an absolutem Ethanol bei -70°C gefällt. Nach einem Waschschnitt mit 70 % Ethanol wurde die DNA getrocknet und anschließend in 50 ml H₂O + RNase (50 mg/ml 30 Endkonzentration) aufgenommen. Die DNA wurde über Nacht bei 4°C gelöst und die RNase-Spaltung wurde anschließend 1 Std. bei 37°C durchgeführt. Die Aufbewahrung der DNA erfolgte bei 4°C.

Beispiel 3: Isolierung von Gesamt-RNA und poly(A)⁺-RNA aus Pflanzen und *Phaeodactylum tricornutum*

40 Die Isolierung von Gesamt-RNA aus Pflanzen wie Lein und Raps etc. erfolgt nach einer bei Logemann et al beschriebenen Methode (1987, Anal. Biochem. 163, 21) isoliert. Aus Moos kann 45 die Gesamt-RNA Protonema-Gewebe nach dem GTC-Verfahren (Reski et al., 1994, Mol. Gen. Genet., 244:352-359) gewonnen werden.

RNA Isolierung aus *Phaeodactylum tricornutum*:

Tiefgefrorene Algenproben (- 70°C) wurden in einem eiskaltem Mörser unter Flüssigstickstoff zu feinem Pulver zerreiben.

5 2 Volumen Homogenisationsmedium (12,024 g Sorbitol, 40,0 ml 1M Tris-HCl, pH 9 (0,2 M); 12,0 ml 5 M NaCl (0,3 M), 8,0 ml 250 mM EDTA, 761,0 mg EGTA, 40,0 ml 10 % SDS wurden auf 200 ml mit H₂O aufgefüllt und der pH auf 8,5 eingestellt) und 4 Volumen Phenol mit 0,2 % Mercaptoethanol wurden bei 40 bis 50°C unter gutem

10 Mischen zu gefrorenem Zellpulver gegeben. Danach wurden 2 Volumen Chloroform hinzufügen und für 15 min kräftig gerührt. Es wurde 10 min bei 10000 g zentrifugiert und die wässrige Phase mit Phenol/Chloroform (2 Vol) und abschließend mit Chloroform extrahiert.

15 Das erhaltene Volumen der wässrigen Phase wurde mit 1/20 Vol 4 M Na-Acetat (pH 6) und 1 Vol Isopropanol (eiskalt) versetzt und die Nukleinsäuren bei -20°C über Nacht (= ÜN) gefällt. Anschließend wurde 30 min bei 10000 g zentrifugiert und der

20 Überstand abgesogen. Es folgte ein Waschschritt mit 70 % EtOH und erneute Zentrifugation. Das Sediment wurde in Tris-Borat-Puffer (80 mM Tris-Borat-Puffer, 10 mM EDTA, pH 7,0) aufgenommen. Dann wurde der Überstand mit 1/3 Vol 8 M LiCl versetzt, gemischt 30 min bei 4°C inkubiert. Nach erneutem zentrifugieren wurde

25 das Sediment mit 70 % Ethanol gewaschen, zentrifugiert und das Sediment in RNase-freiem Wasser gelöst.

Die Isolierung von poly(A)⁺-RNA erfolgte unter Verwendung von Dyna Beads® (Dynal, Oslo, Finnland) nach den Anweisungen im Protokoll

30 des Herstellers.

Nach der Bestimmung der RNA- oder poly(A)⁺-RNA-Konzentration wurde die RNA durch Zugabe von 1/10 Volumina 3 M Natriumacetat, pH 4,6, und 2 Volumina Ethanol gefällt und bei -70°C aufbewahrt.

35 Für die Analyse wurden jeweils 20 µg RNA in einem Formaldehydhaltigen 1,5%igen Agarosegel aufgetrennt und auf Nylon Membranen (Hybond, Amersham) überführt. Der Nachweis spezifischer Transkripte wurde wie bei Amasino beschrieben durchgeführt

40 ((1986) Anal. Biochem. 152, 304)).

Beispiel 4: Konstruktion der cDNA-Bank

Zur Konstruktion der cDNA-Bank aus *Phaeodactylum tricornutum*

45 wurde die Erststrangsynthese unter Verwendung von Reverser Transkriptase aus Maus-Leukämie-Virus (Roche, Mannheim, Deutschland) und Oligo-d(T)-Primern, die Zweitstrangsynthese

durch Inkubation mit DNA-Polymerase I, Klenow-Enzym und RNase H-Spaltung bei 12°C (2 Std.), 16°C (1 Std.) und 22°C (1 Std.) erzielt. Die Reaktion wurde durch Inkubation bei 65°C (10 min) gestoppt und anschließend auf Eis überführt. Doppelsträngige DNA-
5 Moleküle wurde mit T4-DNA-Polymerase (Roche, Mannheim) bei 37°C (30 min) mit glatten Enden versehen. Die Nukleotide wurden durch Phenol/Chloroform-Extraktion und Sephadex-G50-Zentrifugiersäulen entfernt. EcoRI/XhoI-Adapter (Pharmacia, Freiburg, Deutschland) wurden mittels T4-DNA-Ligase (Roche, 12°C, über Nacht) an die
10 cDNA-Enden ligiert, mit XhoI nachgeschnitten und durch Inkubation mit Polynukleotidkinase (Roche, 37°C, 30 min) phosphoryliert. Dieses Gemisch wurde der Trennung auf einem Low-Melting-Agarose-Gel unterworfen. DNA-Moleküle über 300 Basenpaaren wurden aus dem Gel eluiert, Phenol-extrahiert, auf Elutip-D-Säulen (Schleicher
15 und Schüll, Dassel, Deutschland) konzentriert und an Vektorarme ligiert und in lambda-ZAP-Express-Phagen unter Verwendung des Gigapack Gold-Kits (Stratagene, Amsterdam, Niederlande) verpackt, wobei Material des Herstellers verwendet und seine Anweisungen befolgt wurden.

20

Beispiel 5: DNA-Sequenzierung und Computeranalyse

cDNA-Banken, wie im Beispiel 4 beschrieben, wurden zur DNA-Sequenzierung nach Standardverfahren, insbesondere durch das
25 Kettenterminationsverfahren unter Verwendung des ABI PRISM Big Dye Terminator Cycle Sequencing Ready Reaction-Kit (Perkin-Elmer, Weiterstadt, Deutschland), verwendet. Die Sequenzierung zufälliger, vereinzelter Klone wurde anschließend an die präparative Plasmidgewinnung aus cDNA-Banken über in vivo-Massen-
30 excision und Retransformation von DH10B auf Agarplatten durchgeführt (Einzelheiten zu Material und Protokoll von Stratagene, Amsterdam, Niederlande). Plasmid-DNA wurde aus über Nacht gezüchteten *E. coli*-Kulturen, die in Luria-Brühe mit Ampicillin (siehe Sambrook et al. (1989) (Cold Spring Harbor Laboratory
35 Press: ISBN 0-87969-309-6)) gezüchtet worden waren, an einem Qiagen-DNA-Präparations-Roboter (Qiagen, Hilden) nach den Protokollen des Herstellers präpariert. Sequenzierprimer mit den folgenden Nukleotidsequenzen wurden verwendet:

40 5'-CAGGAAACAGCTATGACC-3'
5'-CTAAAGGGAACAAAGCTG-3'
5'-TGTAAAACGACGCCAGT-3'

Die Sequenzen wurden unter Verwendung des Standard-Softwarepaketes EST-MAX, das kommerziell von Bio-Max (München, Deutschland) geliefert wird, prozessiert und annotiert. Durch Nutzung von Vergleichsalgorithmen und unter Verwendung der in SEQ ID NO: 8

dargestellten Suchsequenz wurde mithilfe des BLAST-Programms nach homologen Genen gesucht (Altschul et al. (1997) "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Res. 25:3389-3402.). Zwei Sequenzen aus Phaeodactylum tricornutum mit Homologien zur Suchsequenz aus Physcomitrella patens wurden eingehender charakterisiert.

Beispiel 5a: Isolation von Desaturasen aus Phaeodactylum tricornutum über Polymerase Kettenreaktion mithilfe
10 degenerierter Oligonukleotide:

Mithilfe von publizierten Desaturasen können Motive identifiziert werden, die für Δ-5 und Δ-6 Desaturasen typisch sind. Im folgenden sind Oligonukleotidsequenzen mit möglichen Variationen dargestellt. Unter der Oligonukleotidsequenz ist im Ein-Buchstaben-
15 code die Aminosäure dargestellt, von der die Basenkombination abgeleitet werden kann. Z.B. bedeutet A/G, daß an dieser Position bei der Synthese des Bausteins statistisch gleichverteilt entweder ein A oder ein G in das Oligonukleotid eingebaut wird,
20 da das von der korrespondierenden Aminosäure abgeleitete Basentriplett entweder ein AAA oder ein AAG sein kann. Die DNA Sequenz kann auch ein Inosin (i) enthalten, wenn die Bestimmung einer Base an dieser Position aufgrund des genetischen Codes drei oder vier unterschiedliche Basen erlaubt. Folgende Sequenzen
25 und Primer können verwendet werden:

5'-Vorwärts-Primer:

F1a:	TGG TGG AA A/G TGG	AAi	CA	T/C	AA
F1b:	TGG TGG AA A/G TGG	ACi	CA	T/C	AA
30 F1a:	W W K W	N/T	H		K/N
F1b:	W W K W	K	H		K/N
F2a:	Gi TGG AA A/G GAI	A/C Ai	CA	T/C	AA
F2b:	Gi TGG AA A/G TTG	A/C Ai	CA	T/C	AA
35 F2a:	G/W W K E/D	K/Q/N	H		K/N
F2b:	G/W W K W	K/Q/N	H		K/N
F3a:	T A/T i TTG	AAi	A/C A A/G	C/A G/A i	CA
F3b:	T A/T i TTG	AAi	A/C A A/G	CAi	CA
F3a:	W W	K/N H/N	R/Q		H
40 F3b:	Y W	K/N H/N	R/Q		H
F4a:	GTi TGG A A/T G/A	GA A/G	CA	A/G	CA
F4b:	GTi TGG A A/T G/A	A/T A T/C	CA	A/G	CA
F4a:	V W K/M	E	Q		H
F4b:	V W K/M	N/Y	Q		H
45 F5a1:	CA T/C TA T/C TGG AA A/G AA T/C CA G	C			
F5a1:	CA T/C TA T/C TGG AA A/G AA T/C CA A	C			
F5a1:	H Y W K N Q		H/Q		

80

F6a:	TTG	TTG	AAi	A/C A A/G	AA i	CA T/C	AA
F6a:	W	W	K/N	H/N	K/N	H	K/N

3'- Reverse Primer

5 R1b:	GG	A/G AA	iAG	G/A TG	G/A TG	T/C TC	
R1b:	GG	A/G AA	iAA	G/A TG	G/A TG	T/C TC	
R1a:	P	F	L	H	H	E	
R1b:	P	F	F	H	H	E	
R2a1:	AA	iAG	A/G TG	A/G TG	iA C/T	iA/G T/C TG	
10 R2a2:	AA	T/C AA	A/G TG	A/G TG	iA C/T	iA/G T/C TG	
R2a1:	F	L	H	H	V/I	V/A Q	
R3a1:	AT	iTG	iGG	A/G AA	iAA	A/G TG	A/G TG
R3a2:	AT	A/G TT	iGG	A/G AA	iAA	A/G TG	A/G TG
R3a3:	AT	iTG	iGG	A/G AA	iAG	A/G TG	A/G TG
15 R3a4:	AT	A/G TT	iGG	A/G AA	iAG	A/G TG	A/G TG
R3a1:	I/M H/Q	P	F	F	H	H	
R3a2:	I/M N	P	F	L	H	H	
R4a1:	CT	iGG	A/G AA	iA A/G	A/G TG	A/G TG	
R4a2:	GA	iGG	A/G AA	iA A/G	A/G TG	A/G TG	
20 R4a3:	GT	iGG	A/G AA	iA A/G	A/G TG	A/G TG	
R4a1:	= T/R/S	P	F	F/L	H	H	
R5a1:	AA iAA	A/G TG	A/G TG	T/C TC	T/A/G AT	T/C TG	
R5a2:	AA iAG	A/G TG	A/G TG	T/C TC	T/A/G AT	T/C TG	
R5a1:	F F		H	H	E I	Q	
25 R5a2:	F L		H	H	E I	Q	
R6a1:	T	iGG	iA A/G	iAA	A/G TG	A/G TG	iAC
R6a1:	T	iGG	iA A/G	iAG	A/G TG	A/G TG	iAC
R6a1:	T/N	P	L	F/L	H	H	V

30 Aufgrund verschiedener Variationsmöglichkeiten sind viele abgeleitete Oligonukleotide möglich, jedoch überraschenderweise gefunden wurde, dass dargestellte Oligonukleotide besonders zur Isolation von Desaturasen geeignet sein können.

35 Die Primer können in allen Kombinationen für Polymerase Kettenreaktionen eingesetzt werden. Mithilfe einzelner Kombinationen konnten Desaturase-Fragmente isoliert, wenn nachfolgende Bedingungen berücksichtigt wurden: Für PCR Reaktionen wurden jeweils 10 nMol Primer und 10 ng einer durch in vivo Excision gewonnenen Plasmidbank eingesetzt. Die Plasmidbank konnte nach Protokollen des Herstellers (Stratagene) aus der Phagenbank isoliert werden. Die PCR-Reaktion wurde in einem Thermocycler (Biometra) mit der Pfu-DNA-Polymerase (Stratagene) und dem folgenden Temperaturprogramm durchgeführt: 3 min bei 96°C, gefolgt von 35 Zyklen mit 30 s bei 96°C, 30 s bei 55°C und 1 min bei 72°C. Dabei wurde die Anlagerungstemperatur nach dem ersten Schritt von 55°C schrittweise um je 3°C erniedrigt und nach dem fünften Zyklus

eine Anlagerungstemperatur von 40°C beibehalten. Letztlich wurde ein Zyklus mit 10 min bei 72°C durchgeführt und der Ansatz durch Kühlen auf 4°C beendet.

5 Die Primerkombination F6a und R4a2 sind im Text unterstrichen gekennzeichnet und konnten erfolgreich zur Isolierung eines Desaturasefragmente genutzt werden. Das resultierende Fragment konnte durch Sequenzierung verifiziert werden und zeigte Homologien zu einer Desaturase mit der Genbank Accession Nr. T36617
10 aus Streptomyces coelicolor. Die Homologie wurde mithilfe des BLASTP Programmes erhalten. Der Vergleich ist in Figur 4 dargestellt. Es ergaben sich Identitäten von 34 % und eine Homologie von 43 % zu Sequenz T36617. Das DNA-Fragment wurde gemäß Beispiel 7 in einem Hybridisierungsexperiment zur Isolierung
15 eines Vollängengens nach Standardbedingungen erfindungsgemäß eingesetzt.

Die Codierregion einer so isolierten DNA-Sequenz wurde durch Übersetzung des genetischen Codes in eine Polypeptidsequenz erhalten. In SEQ ID NO: 3 ist eine 1434 Basenpaare lange Sequenz dargestellt, die durch beschriebenes Verfahren isoliert werden konnte. Die Sequenz besitzt ein Startcodon in Position 1 bis 3 und ein Stopcodon in Position 1432-1434 und konnte in ein 477 Aminosäuren langes Polypeptid übersetzt werden. Durch Vergleich mit einer in WO 98 46763 beschriebenen Gensequenz wurde gefunden, dass ein nicht identisches aber homologes Fragment aus Phaeodactylum tricornutum codierend für 87 Aminosäuren vorbeschrieben wurde. Jedoch offenbart WO 98/46763 weder eine vollständige, funktionell aktive Desaturase noch Positions-
25 oder Substratspezifität. Dies wird auch dadurch deutlich, dass sowohl Homologien zur Δ-5, als auch zur Δ-6-Desaturase aus Mortierella alpina berichtet werden, ohne eine genaue Funktion festzulegen. Die erfindungsgemäße Sequenz hingegen codiert für
30 eine funktionell aktive Δ-6-Acyl Lipid Desaturase.
35

Beispiel 6: Identifizierung von DNA Sequenzen codierend für Desaturasen aus Phaeodactylum tricornutum

Die Vollängensequenz der Δ-6-Acyl Lipid Desaturase Pp_des6
40 AJ222980 (NCBI Genbank Accession Nr.) aus dem Moos Physcomitrella patens (siehe auch Tabelle 1) sowie die Δ-12-acyl Lipid Desaturase Sequenz (Tabelle 1 siehe Ma_des12) aus Mortierella alpina AF110509 (AF110509 NCBI Genbank Accession Nr.) wurden für Sequenzvergleiche mithilfe des TBLASTN Suchalgorhythmus
45 eingesetzt.

Die EST-Sequenzen PT0010070010R, PT001072031R sowie PT001078032R wurden zunächst aufgrund schwacher Homologien mit den Suchsequenzen aus *Physcomitrella* und *Mortierella* unter weiteren Kandidatengenen als Zielgen in Betracht gezogen. In Figur 1 und 5 in Figur 2 sowie Figur 2a ist das Ergebnis der zwei gefundenen est-Sequenzen dargestellt. Die gefundenen Sequenzen sind Teil der erfindungsgemäßen Nukleinsäuren aus SEQ ID NO: 1 (Genname: Pt_des5, eigene Datenbank Nr. der Erfinder PT001078032R), SEQ ID NO: 5. (Genname: Pt_des12, eigene Datenbank NR. der 10 Erfinder PT0010070010R) und SEQ ID NO: 11 (Genname: Pt_des12.2, eigene Datenbank des Erfinders PT001072031R). Buchstaben zeigen identische Aminosäuren an, während das Pluszeichen eine chemisch ähnliche Aminosäure bedeutet. Die Identitäten bzw. Homologien aller erfindungsgemäß gefundener Sequenzen gehen aus Tabelle 2 15 zusammenfassend hervor.

Desaturasen können Cytochrom b5 Domänen aufweisen, die auch in anderen nicht Desaturasen codierenden Genen vorkommen. Cytochrom b5 Domänen zeigen mithin hohe Homologien an, obwohl es sich um 20 verschiedene Genfunktionen handelt. Desaturasen können schwach konservierter Bereiche lediglich als putative Kandidatengene identifiziert werden und müssen auf die Enzymaktivität und Positionsspezifität der enzymatischen Funktion hin geprüft werden. Beispielsweise zeigen auch verschiedene Hydroxylasen, 25 Acetylenasen und Epoxygenasen ähnlich wie Desaturasen Histidin-Box Motive, so dass eine konkrete Funktion experimentell nachgewiesen werden muß und zusätzlich die Verifizierung der Doppelbindung erst eine sichere Enzymaktivität und Positionsspezifität einer Desaturase ermöglicht. Überraschenderweise wurde gefunden, 30 dass erfindungsgemäße Δ-6- und Δ-5- Desaturase besonders geeignete Substratspezifitäten aufweisen und besonders geeignet sind, um in Kombination mit einer Δ-6-Elongase aus *Physcomitrella* zur Produktion von polyungesättigten Fettsäuren wie Arachidonsäure, Eicosapentaensäure und Docosahexaensäure genutzt werden können.

35 Die Sequenzierung des vollständigen cDNA Fragmentes aus Klon PT001078032R ergab eine 1652 Basenpaare lange Sequenz. Die Sequenz codiert für ein Polypeptid von 469 Aminosäuren dargestellt in SEQ ID NO: 2. Diese wurde erhalten durch Über- 40 setzung des genetischen Codes aus SEQ ID NO: 1 mit einem Startcodon in Basenpaarposition 115-117 und mit einem Stopcodon in Basenpaarposition 1522-1524. Der Klon beinhaltet ein vollständiges Desaturase-Polypeptid, wie aus dem Sequenzvergleich in Figur 3 zu ersehen ist. Striche bedeuten identische Amino- 45 säuren während Doppelpunkte und Einzelpunkte chemisch austauschbare, d.h. chemisch äquivalente Aminosäuren darstellen. Der Vergleich wurde mit der BLOSUM62 Austauschmatrix für Amino-

83

säuren nach Henikoff & Henikoff durchgeführt: ((1992) Amino acid substitution matrices from protein blocks. Proc. Natl. Acad. Sci. USA 89: 10915-10919). Verwendete Parameter: Gap Weight: 8; Average Match: 2.912, Length Weight: 2, Average Mismatch: -2.003.

5

In Figur 6 und Figur 7 ist der Vergleich der MA_des12 Peptidsequenz mit den gefundenen Sequenzen dargestellt.

Die Sequenzierung des vollständigen cDNA Fragmentes aus Klon 10 PT0010070010R ergab eine in SEQ ID NO: 5 dargestellte 1651 Basenpaare lange Sequenz mit einem Startcodon in Position 67-69 und einem Stopcodon in Position 1552-1554. Die erfindungsgemäße Polypeptidsequenz ist in SEQ ID NO: 6 dargestellt.

15 Die Sequenzierung des vollständigen identifizierten cDNA Fragmentes aus Klon PT0010072031R ergab eine in SEQ ID NO: 11 dargestellte 1526 Basenpaare lange Sequenz mit einem Startcodon in Position 92-94 und einem Stopcodon in Position 1400-1402. Die erfindungsgemäße Polypeptidsequenz ist in SEQ ID NO: 12 20 dargestellt.

In Tabelle 2 sind die Identitäten und Homologien erfindungsgemäßer Desaturasen untereinander und mit der Desaturase aus *Physcomitrella patens* und *Mortierella alpina* dargestellt. Die 25 Angaben wurden mithilfe des Programms Bestfit unter gegebenen Parametern wie unten definiert als Teilprogramm folgender Software erhalten: Wisconsin Package Version 10.0 (Genetics Computer Group (GCG), Madison, Wisc., USA). Henikoff, S. and Henikoff, J.G. (1992). Amino acid substitution matrices from protein 30 blocks. Proc. Natl. Acad. Sci. USA 89: 10915-10919.

Weiterhin ist in Figur 5 der Vergleich der Δ-6-acyl Lipid Desaturase aus *Physcomitrella patens* mit der Polypeptidsequenz des Klons Pt_des6 dargestellt.

35

Tabelle 2:

	Homologie / Identität in %	Suchsequenz Pp_des6	Suchsequenz Ma_des12
40	Pt_des5	34.92/26.37	n.d.
	Pt_des6	50.69/41.06	n.d.
	Pt_des12	n.d.	48.58/38.92
	Pt_des12.2	n.d.	48.37/41.60

45 n.d. = nicht durchgeführt

Mithilfe des Algorhythmus TBLASTN 2.0.10: Altschul et al 1997, "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Res. 25:3389-3402 wurden über einen lokalen Datenbankvergleich Sequenzen mit höchster Sequenz-
5 homologie bzw. Identität identifiziert. Die Ergebnisse sind in folgender Tabelle 2A dargestellt.

Tabelle 2A: Homologe mit den höchsten Sequenzhomologien bzw
Identitäten zu erfindungsgemäßen Polypeptidsequenzen
10 aus SEQ ID NO. 2, 4, 6 oder 12

	Homologie / Identität (%)	Suchsequenz PT001070010R	Suchsequenz PT001072031R	Suchsequenz PT001078032R	Suchsequenz Pt_des6
15	L26296: Fad2 A. thaliana	50 % / 37 %	n.d.	n.d.	n.d.
	U86072 Petro- selinum crispum Fad2	n.d.	51/40	n.d.	n.d.
20	AL358652 L. major putative desaturase	n.d.	n.d.	45/30	n.d.
25	AB020032 M. alpina delta 6 desaturase	n.d.	n.d.	n.d.	53/38

Beispiel 7: Identifikation von Genen mittels Hybridisierung

Gensequenzen lassen sich zur Identifikation homologer oder
30 heterologer Gene aus cDNA- oder genomischen Banken verwenden.

Homologe Gene (d.h. Voll-Längen-cDNA-Klone, die homolog sind, oder Homologen) lassen sich über Nukleinsäurehybridisierung unter Verwendung von beispielsweise cDNA-Banken isolieren: Ins-
35 besondere zur Isolierung von funktionell aktiven Voll-Längengenen der in SEQ ID NO: 3 gezeigten kann die Methode genutzt werden. Je nach der Häufigkeit des Gens von Interesse werden 100000 bis zu 1000000 rekombinante Bakteriophagen plattiert und auf eine Nylonmembran überführt. Nach der Denaturierung mit Alkali wurde
40 die DNA auf der Membran z.B. durch UV-Vernetzung immobilisiert. Die Hybridisierung erfolgt bei hoch-stringenten Bedingungen. In wässriger Lösung werden die Hybridisierung und die Waschschrifte bei einer Ionenstärke von 1 M NaCl und einer Temperatur von 68°C durchgeführt. Hybridisierungssonden wurden z.B. durch Markierung
45 mittels radioaktiver (³²P-) Nicktranskription (High Prime, Roche,

Mannheim, Deutschland) hergestellt. Die Signale werden mittels Autoradiographie nachgewiesen.

Partiell homologe oder heterologe Gene, die verwandt, aber nicht 5 identisch sind, lassen sich analog zum oben beschriebenen Verfahren unter Verwendung niedrig-stringenter Hybridisierungs- und Waschbedingungen identifizieren. Für die wässrige Hybridisierung wurde die Ionenstärke gewöhnlich bei 1 M NaCl gehalten, wobei die Temperatur nach und nach von 68 auf 42°C gesenkt wurde.

10

Die Isolierung von Gensequenzen, die nur zu einer einzelnen Domäne von beispielsweise 10 bis 20 Aminosäuren Homologien aufweisen, lässt sich unter Verwendung synthetischer, radioaktiv markierter Oligonukleotidsonden durchführen. Radioaktiv markierte 15 Oligonukleotide werden mittels Phosphorylierung des 5'-Endes zweier komplementärer Oligonukleotide mit T4-Polynukleotidkinase hergestellt. Die komplementären Oligonukleotide werden aneinander hybridisiert und ligiert, so dass Konkatenere entstehen. Die doppelsträngigen Konkatenere werden beispielsweise durch Nick- 20 transkription radioaktiv markiert. Die Hybridisierung erfolgt gewöhnlich bei niedrig-stringenten Bedingungen unter Verwendung hoher Oligonukleotidkonzentrationen.

Oligonukleotid-Hybridisierungslösung:

25

6 x SSC
0,01 M Natriumphosphat
1 mM EDTA (pH 8)
0,5 % SDS
30 100 mikrog/ml denaturierte Lachssperma-DNA
0,1 % fettaime Trockenmilch

Während der Hybridisierung wird die Temperatur schrittweise auf 5 bis 10°C unter die berechnete Oligonukleotid-Tm oder bis auf Raum- 35 temperatur (bedeutet RT = ~ 23°C in allen Experimenten, wenn nicht anders angegeben) gesenkt, gefolgt von Waschschriften und Autoradiographie. Das Waschen wird mit extrem niedriger Stringenz durchgeführt, zum Beispiel 3 Waschschriften unter Verwendung von 4 X SSC. Weitere Einzelheiten sind wie von Sambrook, J., et al. 40 (1989), "Molecular Cloning: A Laboratory Manual", Cold Spring Harbor Laboratory Press, oder Ausubel, F.M., et al. (1994) "Current Protocols in Molecular Biology", John Wiley & Sons, beschrieben.

45

Beispiel 8: Identifikation von Zielgenen durch Sichtung von Expressionsbanken mit Antikörpern

Es wurden cDNA-Sequenzen zur Herstellung von rekombinantem Protein zum Beispiel in *E. coli* verwendet (z.B. Qiagen QIAexpress PQE-System). Die rekombinanten Proteine wurden dann gewöhnlich über Ni-NTA-Affinitätschromatographie (Qiagen) affinitätsgereinigt. Die rekombinanten Proteine wurden dann zur Herstellung spezifischer Antikörper beispielsweise unter Verwendung von Standardtechniken zur Immunisierung von Kaninchen verwendet.

Anschließend wurden die Antikörper dann unter Verwendung einer Ni-NTA-Säule, die mit rekombinantem Antigen vorgesättigt wird, affinitätsgereinigt, wie von Gu et al., (1994) BioTechniques 17:257-262 beschrieben. Der Antikörper kann dann zur Durchmusterung von Expressions-cDNA-Banken mittels immunologischem Sichtung verwendet werden (Sambrook, J., et al. (1989), "Molecular Cloning: A Laboratory Manual", Cold Spring Harbor Laboratory Press, oder Ausubel, F.M., et al. (1994) "Current Protocols in Molecular Biology", John Wiley & Sons).

20 Beispiel 9: Transformation von Agrobacterium

Die Agrobacterium-vermittelte Pflanzentransformation kann zum Beispiel unter Verwendung des GV3101- (pMP90-) (Koncz und Schell, 25 Mol. Genet. 204 (1986) 383-396) oder LBA4404- (Clontech) oder C58C1 pGV2260 (Deblaere et al 1984, Nucl. Acids Res. 13, 4777-4788) Agrobacterium tumefaciens-Stamms durchgeführt werden. Die Transformation kann durch Standard-Transformationstechniken durchgeführt werden (ebenfalls Deblaere et al. 1984).

30 Beispiel 10: Pflanzentransformation

Die Agrobacterium-vermittelte Pflanzentransformation kann unter Verwendung von Standard-Transformations- und Regenerations-techniken durchgeführt werden (Gelvin, Stanton B., Schilperoort, Robert A., Plant Molecular Biology Manual, 2. Aufl., Dordrecht: Kluwer Academic Publ., 1995, in Sect., Ringbuc Zentrale Signatur: BT11-P ISBN 0-7923-2731-4; Glick, Bernard R., Thompson, John E., Methods in Plant Molecular Biology and Biotechnology, Boca Raton: 40 CRC Press, 1993, 360 S., ISBN 0-8493-5164-2).

Beispielsweise kann Raps mittels Kotyledonen- oder Hypokotyl-transformation transformiert werden (Moloney et al., Plant Cell 8 (1989) 238-242; De Block et al., Plant Physiol. 91 (1989) 45 694-701). Die Verwendung von Antibiotika für die Agrobacterium- und Pflanzenselektion hängt von dem für die Transformation verwendeten binären Vektor und Agrobacterium-Stamm ab. Die

Rapsselektion wird gewöhnlich unter Verwendung von Kanamycin als selektierbarem Pflanzenmarker durchgeführt.

Der Agrobacterium-vermittelte Gentransfer in Lein (*Linum usitatissimum*) lässt sich unter Verwendung von beispielsweise einer von Mlynarova et al. (1994) Plant Cell Report 13:282-285 beschriebenen Technik durchführen.

Die Transformation von Soja kann unter Verwendung von beispielsweise einer in EP-A-0 0424 047 (Pioneer Hi-Bred International) oder in EP-A-0 0397 687, US 5,376,543, US 5,169,770 (University Toledo) beschriebenen Technik durchgeführt werden.

Die Pflanzentransformation unter Verwendung von Teilchenbeschuss, Polyethylenglycol-vermittelter DNA-Aufnahme oder über die Siliziumcarbonatfaser-Technik ist beispielsweise beschrieben von Freeling und Walbot "The maize handbook" (1993) ISBN 3-540-97826-7, Springer Verlag New York).

20 Beispiel 11: Plasmide für die Pflanzentransformation

Geeignete binäre Vektoren und Transformationsmarker

Zur Pflanzentransformation können binäre Vektoren, wie pBinAR (Höfgen und Willmitzer, Plant Science 66 (1990) 221-230) oder pGPTV (Becker et al. 1992, Plant Mol. Biol. 20:1195-1197) verwendet werden. Die Konstruktion der binären Vektoren kann durch Ligation der cDNA in Sense- oder Antisense-Orientierung in T-DNA erfolgen. 5' der cDNA aktiviert ein Pflanzenpromotor die Transkription der cDNA. Eine Polyadenylierungssequenz befindet sich 3' von der cDNA. Die binären Vektoren können unterschiedliche Markergene tragen. So kann etwa die Resistenz durch die Expression des nptII Markergens unter Kontrolle des 35S oder des nos Promoters erfolgen. Insbesondere kann das nptII-Markergen codierend für Kanamycin-Resistenz vermittelt durch Neomycinphosphotransferase gegen die herbizidresistente Form eines Acetolactat Synthasegens (AHAS oder ALS) ausgetauscht werden. Das ALS-Gen ist beschrieben in Ott et al., J. Mol. Biol. 1996, 263:359-360. Geeignet für manche Pflanzen ist auch die Verwendung des Hygromycin-Resistenz-Gens. Der v-ATPase-c1-Promotor kann in das Plasmid 35 transferiert werden und durch Klonierung vor die kodierende Region des ALS Gens für die Markergenexpression genutzt werden. Der genannte v-ATPase-c1-Promotor entspricht einem 1153 Basenpaarfragment aus Beta vulgaris (Plant Mol Biol, 1999, 39:463-475). Der genannte nos Promoter Dabei können sowohl Sulfonylharnstoffe als auch Imidazolinone wie Imazethapyr oder Sulfonylharnstoffe als Antimetaboliten zur Selektion verwendet werden.

den. Alternativ kann auch der nos-Promoter für die Markergenexpression verwendet werden.

Beispiel 12: Ermittlung von geeigneten Promotoren für die
5 Expression in Lein

Die gewebespezifische Expression lässt sich unter Verwendung eines gewebespezifischen Promotors erzielen. Beispielsweise kann die samenspezifische Expression erreicht werden, indem der DC3-
10 oder der LeB4- oder der USP-Promotor oder der Phaseolin-Promotor 5' der cDNA einkloniert wird. Auch jedes andere samenspezifische Promotorelement wie z.B. der Napin- oder Arcelin Promotor (Goossens et al. 1999, Plant Phys. 120(4):1095-1103 und Gerhardt et al. 2000, Biochimica et Biophysica Acta 1490(1-2):87-98) kann
15 verwendet werden. Zur konstitutiven Expression in der ganzen Pflanzen lässt sich der CaMV-35S-Promotor oder ein v-ATPase-c1 Promotor verwenden.

Um die Eigenschaften des Promotors zu bestimmen und die essen-
20 tiellen Elemente desselben, die seine Gewebespezifität ausmachen, zu identifizieren, ist es erforderlich, den Promotor selbst oder verschiedene Fragmente desselben vor ein sogenanntes Reportergen zu setzen, das eine Bestimmung der Expressionsaktivität ermöglicht. Beispielhaft für ein Reportergen sei die bakterielle
25 β -Glucuronidase (GUS) genannt (Jefferson et al., EMBO J. 1987, 6, 3901-3907). Die β -Glucuronidase Aktivität kann *in-situ* mittels eines chromogenen Substrates wie 5-Bromo-4-Chloro-3-Indolyl- β -D-Glucuronsäure im Rahmen einer Aktivitätsfärbung bestimmt werden (Jefferson, 1987, Plant Molecular Biology Reporter 5,
30 387-405). Für die Untersuchungen der Gewebespezifität wird das pflanzliche Gewebe geschnitten, eingebettet, gefärbt und analysiert wie beschrieben (z.B. Bäumlein H et al., 1991 Mol Gen Genet 225: 121-128).

35 Fluorimetrischer GUS-Test (nach Montgomery et al., 1993)

Dieser Assay erlaubt eine quantitative Bestimmung der GUS-Aktivität in dem untersuchten Gewebe. Für die quantitative Aktivitätsbestimmung wird als Substrat für die β -Glucuronidase MUG (4-Methyl-
40 umbelliferyl-beta-D-glucuronid) verwendet, das in MU (Methyl-umbelliferon) und Glucuronsäure gespalten wird.

Dabei wird zunächst ein Proteinextrakt des gewünschten Gewebe hergestellt, dem dann das Substrat der GUS zugesetzt wird. Das
45 Substrat ist erst nach der Umsetzung durch GUS fluorimetrisch messbar. Zu verschiedenen Zeitpunkten werden Proben entnommen, die anschließend im Fluorimeter gemessen werden. Dieser Test

wurde mit Leinembryonen verschiedener Altersstadien durchgeführt (21, 24 oder 30 Tage nach Beginn der Blüte, daf = days after flowering). Dazu wurde je ein Embryo in einem 2 ml-Reaktionsgefäß mit Hilfe einer Schwingmühle (Retsch MM 2000) in flüssigem Stickstoff zu Pulver zerrieben. Nach Zugabe von 100 ml EGL-Puffer wurde für 10 min bei 25°C und 14000 x g zentrifugiert. Der Überstand wurde abgenommen und ein zweites Mal zentrifugiert. Wieder wurde der Überstand in ein neues Reaktionsgefäß überführt und bis zur weiteren Verwendung auf Eis gehalten. Von diesem Proteinextrakt wurden 25 ml mit 65 ml EGL-Puffer (ohne DTT) versetzt und für den GUS-Assay eingesetzt. Nun wurden 10 ml des Substrates MUG (10 mM 4-Methyl-umbelliferyl-β-D-glucuronid) dazugegeben, geworfen und sofort 30 ml als Nullwert entnommen und mit 470 ml Stopp-Puffer (0,2 M Na₂CO₃) versetzt. Dieser Vorgang wurde für alle Proben in einem Abstand von 30 s wiederholt. Die entnommenen Proben wurden bis zur Messung im Kühlschrank gelagert. Weitere Messwerte wurden nach 1 h und nach 2 h entnommen. Für die Messung im Fluorimeter wurde eine Eichreihe erstellt, die Konzentrationen von 0,1 mM bis 10 mM MU (4-Methyl-umbelliferon) enthielt. Waren die Probenwerte außerhalb dieser Konzentrationen, wurde weniger Proteinextrakt eingesetzt (10 ml, 1 ml, 1 ml aus 1:10 Verdünnung), und es wurden kürzere Zeitabstände gemessen (0 h, 30 min, 1 h). Die Messung erfolgte bei einer Excitation von 365 nm und einer Emission von 445 nm in einem Fluoroscan II-Gerät (Labsystem). Alternativ kann die Substratsspaltung unter alkalischen Bedingungen fluorometrisch verfolgt werden (Anregung bei 365 nm, Messung der Emission bei 455 nm; SpectroFluorimeter BMG Polarstar+) wie beschrieben in Bustos M.M. et al., 1989 Plant Cell 1:839-853. Alle Proben wurden einer Proteinkonzentrationsbestimmung nach Bradford (1976) unterzogen, um so eine Aussage über die Promoteraktivität und -stärke in verschiedenen Geweben und Pflanzen erlauben.

EGL-Puffer

35 0,1 M	KPO ₄ , pH 7,8
1mM	EDTA
5 %	Glycerin
1 M	DTT

Als weitere Beispiele für Reportergene, die äquivalent benutzt werden können, seien beispielhaft das grün fluoreszierende Protein (GFP) und dessen Derivate genannt (C.Reichel et al. (1996) Proc.Natl.Acad.Sci.USA 93, 5888-5893 und J.Sheen et al., (1995) Plant Journal 8, 777-784) und verschiedene Luciferasen (A.Millar et al. (1992) Plant Mol. Biol. Reporter 10, 324-414). Die ent-

sprechenden Detektionsmethoden sind dem Fachmann bekannt und z.B. genannter Literatur beschrieben.

Beispiele für Promoter-Reportergen-Konstrukte für oben genannte
5 Promotoren sind im folgenden gegeben. Von diesen Promotoren können Fragmente mithilfe der Polymerasekettenreaktion isoliert und mit flankierenden Sequenzen nach Wahl auf Basis von synthetischen Oligonukleotiden maßgeschneidert werden.

10 Folgende Oligonukleotide können beispielsweise verwendet werden:

LeB4 vorne: GAAAGCTTCTCGAGTTATGCATTTCTT
LeB4 hinten: GGGTCTAGATCTGTGACTGTGATAG
DC3a vorne: AGTGGATCCCCGAGCTAACACACA
15 DC3a hinten: ATAAGCTTTCTTTGCAGA
napinvorne: GAAAGCTTCTAATATGATAAACTCTG
napinhinten: GGGTCTAGAACACATACAAACATCAC

Die Methoden sind dem Fachmann auf dem Gebiet bekannt und sind
20 allgemein literaturbekannt.

In einem ersten Schritt werden die Promotorfragmente über PCR amplifiziert, mit geeigneten Restriktionsenzymen geschnitten und in die obigen Kassetten einkloniert. Beispielsweise wird das
25 LeB4(700)-PCR-Fragment mit XbaI und HindIII geschnitten und in den Vektor pGPTV in die HindIII und XbaI-Schnittstellen 5' vor dem GUS Reportergen kloniert. Das PCR-amplifizierte DC3-Promoterfragment kann mit BamHI und HindIII geschnitten, z.B. in pBluescript (Stratagene) subkloniert und dann in pGPTV in geeignete
30 Schnittstellen vor das GUS-Reportergen kloniert werden.

Beispielsweise kann ein mit obigen Primern PCR-amplifiziertes napin-Promotorfragment mit einer Größe von 1055bp nach Verdau mit HindIII und XbaI in pGPTV vor das GUS Reportergen einkloniert
35 werden. Ein äquivalentes Konstrukt könnte ein Napin-Promoterfragment von 1100bp 5'-kloniert vor ein GUS Reportergen mit Intron mit in 3' Richtung nachfolgendem Nos-Terminator in einem pHL9000-Vektor (Hausmann & Töpfer, 1999) sein.

40 Unter Verwendung von oben beschriebenen oder äquivalenten Konstrukten nach Transformation in Lein der Sorte Flanders, lässt sich die GUS-Aktivität in transgenen Leinembryonen verschiedener Altersstadien messen, die mit einem der folgenden Konstrukte transformiert wurden: Napin-GUS, 35S-GUS, LeB4-GUS, USP-GUS. Die
45 Werte sind Mittelwerte aus ein bis fünf Messungen mit verschiede-

nen Proteinmengen. Pro Konstrukt wurden je drei Embryonen quantitativ analysiert.

In der quantitativen Analyse ergaben sich große Unterschiede zwischen den verschiedenen samenspezifischen Promotoren. Der Napin Promotor aus *Brassica napus* erwies sich als um zwei bis drei Zehnerpotenzen weniger aktiv als die beiden Promotoren aus *Vicia faba* (LeB4 und USP). Die GUS-Aktivität nahm in der Reihenfolge Napin, LeB4 und USP zu. Die Positivkontrolle 35S bewegte sich in ihrer Aktivität zwischen LeB4 und USP, wohingegen die Negativkontrolle, der nicht transformierte Wildtyp (Sorte Flanders), so gut wie keine Aktivität besaß. Die folgende Tabelle gibt die Mittelwerte der Aktivitäten in den einzelnen Altersstadien sowie gesamt für jedes Konstrukt wieder.

15

Tabelle 3.4: Übersicht über die mittleren GUS-Aktivitäten von Leinembryonen, transformiert mit verschiedenen GUS-Konstrukten. daf: Tage nach Beginn der Blüte.

20 Mit * markiert wurden Werte, die nur einer Messung zugrundeliegen.

	GUS-Konstrukt	GUS-Aktivität [nmol/h/mg Protein]	GUS-Aktivität [nmol/h/mg Protein]	GUS-Aktivität [nmol/h/mg Protein]	GUS-Aktivität [nmol/h/mg Protein]	% vs. 35S
25		Mittelwert 21 daf	Mittelwert 24 daf	Mittelwert 30 daf	Gesamtmittelwert	
30	Ohne 35S	0,06 649,00	0,06 913,00	*0,02 639,00	0,05 734,00	0,0007 100,00
	Napin	7,70	5,70	3,70	5,70	0,80
	LeB4	1778,00	253,00	283,00	771,00	105,00
35	USP	2843,00	3770,00	2107,00	2907,00	396,00

Beispiel 13: Plasmide für die Pflanzentransformation

Zur Pflanzentransformation können binäre Vektoren, wie pBinAR (Höfgen und Willmitzer, Plant Science 66 (1990) 221-230) oder PGPTV (Becker et al 1992, Plant Mol. Biol. 20:1195-1197) oder Derivate davon verwendet werden. Die Konstruktion der binären Vektoren kann durch Ligation der cDNA in Sense- oder Antisense-Orientierung in T-DNA erfolgen. 5' der cDNA aktiviert ein Pflanzenpromotor die Transkription der cDNA. Eine Polyadenylierungssequenz befindet sich 3' von der cDNA. Die binären Vektoren können unterschiedliche Markergene tragen. Insbesondere kann das nptII-Markergen codierend für Kanamycin-Resistenz vermittelt

92

durch Neomycinphosphotransferase gegen die herbizidresistente Form eines Acetolactat Synthasegens (Abkürzung: AHAS oder ALS) ausgetauscht werden. Das ALS-Gen ist beschrieben in Ott et al., J. Mol. Biol. 1996, 263:359-360. Der v-ATPase-c1-Promotor kann in 5 das Plasmid pBin19 oder pGPTV kloniert werden und durch Klonierung vor das ALS Codierregion für die Markergenexpression genutzt werden. Der genannte Promotor entspricht einem 1153 Basenpaarfragment aus beta-Vulgaris (Plant Mol Biol, 1999, 39:463-475). Dabei können sowohl Sulphonylharnstoffe als auch Imidazolinone 10 wie Imazethapyr oder Sulphonylharnstoffe als Antimetaboliten zur Selektion verwendet werden.

Die gewebespezifische Expression lässt sich unter Verwendung eines gewebespezifischen Promoters erzielen. Beispielsweise kann 15 die samenspezifische Expression erreicht werden, indem der DC3- oder der LeB4- oder der USP-Promotor oder der Phaseolin-Promotor 5' der cDNA einkloniert wird. Auch jedes andere samenspezifische Promotorelement wie z.B. der Napin- oder Arcelin Promotor Goossens et al. 1999, Plant Phys. 120(4):1095-1103 und Gerhardt 20 et al. 2000, Biochimica et Biophysica Acta 1490(1-2):87-98) kann verwendet werden. Zur konstitutiven Expression in der ganzen Pflanzen lässt sich der CaMV-35S-Promotor oder ein v-ATPase C1 Promotor verwenden.

25 Insbesondere lassen sich Gene codierend für Desaturasen und Elongasen durch Konstruktion mehrerer Expressionskassetten hintereinander in einen binären Vektor klonieren, um den Stoffwechselweg in Pflanzen nachzubilden.

30 Innerhalb einer Expressionskassette kann das zu exprimierende Protein unter Verwendung eines Signalpeptids, beispielsweise für Plastiden, Mitochondrien oder das Endoplasmatische Retikulum, in ein zelluläres Kompartiment dirigiert werden (Kermode, Crit. Rev. Plant Sci. 15, 4 (1996) 285-423). Das Signalpeptid wird 5' im 35 Leseraster mit der cDNA einkloniert, um die subzelluläre Lokalisierung des Fusionsprotein zu erreichen.

Beispiele für Multiexpressionskassetten sind im folgenden gegeben.

40

I.) Promotor-Terminator-Kassetten

Expressionskassetten bestehen aus wenigstens zwei funktionellen Einheiten wie einem Promotor und einem Terminator. Zwischen 45 Promotor und Terminator können weitere gewünschte Gensequenzen wie Targetting-Sequenzen, Codierregionen von Genen oder Teilen davon etc. eingefügt werden. Zum Aufbau von Expressionskassetten

werden Promotoren und Terminatoren (USP Promotor: Baeumlein et al., Mol Gen Genet, 1991, 225 (3):459-67); OCS Terminator: Gielen et al. EMBO J. 3 (1984) 835ff.) mithilfe der Polymerasekettenreaktion isoliert und mit flankierenden Sequenzen nach Wahl 5 auf Basis von synthetischen Oligonukleotiden maßgeschneidert.

Folgende Oligonukleotide können beispielsweise verwendet werden:

USP1 vorne: CCGGAATTGGCGCGCCGAGCTCCTCGAGCAAATTTACACATTGCCA

USP2 vorne: CCGGAATTGGCGCGCCGAGCTCCTCGAGCAAATTTACACATTGCCA

10 USP3 vorne: CCGGAATTGGCGCGCCGAGCTCCTCGAGCAAATTTACACATTGCCA

USP1 hinten: AAAACTGCAGGGCGGCCACCGCGGTGGCTGGCTATGAAGAAATT

USP2 hinten: CGCGGATCCGCTGGCTATGAAGAAATT

USP3 hinten: TCCCCCGGGATCGATGCCGGCAGATCTGCTGGCTATGAAGAAATT

OCS1 vorne: AAAACTGCAGTCTAGAAGCCTCTGCTTAATGAGATAT

15 OCS2 vorne: CGCGGATCCGATATGGGCCGCTAGCGTTAACCTGCTTAATGAGATAT

OCS3 vorne: TCCCCCGGGCATGGCTGCTTTAATGAGATAT

OCS1 hinten: CCCAAGCTTGGCGCGCCGAGCTCGAATTCTGCGACGGACAATCAGTAAATTGA

OCS2 hinten: CCCAAGCTTGGCGCGCCGAGCTCGACGGACAATCAGTAAATTGA

OCS3 hinten: CCCAAGCTTGGCGCGCCGAGCTCGACGGACAATCAGTAAATTGA

20

Die Methoden sind dem Fachmann auf dem Gebiet bekannt und sind allgemein literaturbekannt.

In einem ersten Schritt werden ein Promotor und ein Terminator 25 über PCR amplifiziert. Dann wird der Terminator in ein Empfängerplasmid kloniert und in einem zweiten Schritt der Promotor vor den Terminator inseriert. Mithin erhält man eine Expressionskassette auf einem Trägerplasmid. Auf Basis des Plasmides pUC19 werden die Plasmide pUT1, pUT2 und pUT3 erstellt.

30

Die Konstrukte sind erfindungsgemäß in SEQ ID NO: 13, 14 und 15 definiert. Sie enthalten auf Basis von pUC19 den USP-Promotor und den OCS Terminator. Auf Basis dieser Plasmide wird das Konstrukt pUT12 erstellt, indem pUT1 mittels SalI/ScalI geschnitten wird und 35 pUT2 mittels XhoI/ScalI geschnitten wird. Die die Expressionskassetten enthaltenden Fragmente werden ligiert und in E. coli XLI blue MRF transformiert. Es wird nach Vereinzelung ampicillinresistenter Kolonien DNA präpariert und per Restriktionsanalyse solche Klone identifiziert, die zwei Expressionskassetten enthalten. Die XhoI/SalI Ligation kompatibler Enden hat dabei die 40 beiden Schnittstellen XhoI und SalI zwischen den Expressionskassetten eliminiert. Es resultiert das Plasmid pUT12, das in SEQ ID NO: 16 definiert ist. Anschließend wird pUT12 wiederum mittels SalI/ScalI geschnitten und pUT3 mittels XhoI/ScalI geschnitten. Die die Expressionskassetten enthaltenden Fragmente werden ligiert und in E. coli XLI blue MRF transformiert. Es wird 45 nach Vereinzelung ampicillinresistenter Kolonien DNA präpariert

und per Restriktionsanalyse solche Klone identifiziert, die drei Expressionskassetten enthalten. Auf diese Weise wird ein Set von Multiexpressionskassetten geschaffen, dass für die Insertion gewünschter DNA genutzt werden kann und in Tabelle 3 beschrieben 5 wird und zudem noch weitere Expressionskassetten aufnehmen kann.

Diese enthalten folgende Elemente:

Tabelle 3

10

	pUC19-Derivat	Schnittstellen vor dem USP Promotor	Multiple Klonierungs-Schnittstellen	Schnittstellen hinter dem OCS-Terminator
	pUT1	EcoRI/AsclI/ SacI/XhoI	BstXI/NotI/ PstI/XbaI/StuI	SalI/EcoRI/ SacI/AsclI/HindIII
15	pUT2	EcoRI/AsclI/ SacI/XhoI	BamHI/EcoRV/ ApaI/NheI/ HpaI	SalI/EcoRI/ SacI/AsclI/HindIII
	pUT3	EcoRI/AsclI/ SacI/XhoI	BglII/NaeI/ ClaI/SmaI/NcoI	SalI/SacI/ AsclI/HindIII
	pUT12 Doppel-expressionskassette	EcoRI/AsclI/ SacI/XhoI	BstXI/NotI/ PstI/XbaI/StuI Und BamHI/EcoRV/ ApaI/NheI/ HpaI	SalI/EcoRI/ SacI/AsclI/HindIII
20			1.BstXI/NotI/ PstI/XbaI/StuI und 2.BamHI/EcoRV/ ApaI/NheI/ HpaI und 3.BglII/NaeI/ ClaI/SmaI/NcoI	
25	pUT123 Tripel-expressionskassette	EcoRI/AsclI/ SacI/XhoI		SalI/SacI/AsclI/HindIII

Weiterhin lassen sich wie beschrieben und wie in Tabelle 4 näher spezifiziert weitere Multiexpressionskassetten mithilfe des

30 i) USP-Promotors oder mithilfe des
ii) ca. 700 Basenpaare 3'-Fragmentes des LeB4-Promotors oder
mithilfe des
iii) DC3-Promotors erzeugen und für samenspezifische Genexpression
einsetzen.

35

Der DC3-Promotor ist beschrieben bei Thomas, Plant Cell 1996,
263:359-368 und besteht lediglich aus der Region -117 bis +27
weshalb er mithin einer der kleinsten bekannten samenspezifischen
Promotoren darstellt.

40

Von diesen Promotoren können Fragmente mithilfe der Polymerasekettenreaktion isoliert und mit flankierenden Sequenzen nach Wahl auf Basis von synthetischen Oligonukleotiden maßgeschneidert werden.

45

Folgende Oligonukleotide können beispielsweise verwendet werden:

LeB4 vorne: GAAAGCTTCTCGAGTTATGCATTCCTT
LeB4 hinten: GGGTCTAGATCTGTGACTGTGATAG
DC3a vorne: CCGGAATTGGCGCGCCGAGCTCCTCGAG
DC3a hinten: CGCGGATCCTAGCTTTCTTGGCAGATG

5

Die Methoden sind dem Fachmann auf dem Gebiet bekannt und sind allgemein literaturbekannt.

In einem ersten Schritt werden die Promotorfragmente über PCR am-
10 plifiziert, mit geeigneten Restriktionsenzymen geschnitten und in die obigen Kassetten einkloniert. Beispielsweise wird das LeB4(700)-PCR-Fragment mit XhoI und BglII geschnitten und in die XhoI und BglII-Schnittstellen des Plasmids pUT3 eingesetzt, um pLT3 zu erhalten.

15

Vorteilhafte Expressionskassetten enthalten auf Basis von pUC19 (Vieira und Messing (1982); Gene 19, 259), die SEQ ID NO: 32, den LeB4-Promotor und die Sequenzen SEQ ID NO: 33, SEQ ID NO: 34 oder SEQ ID NO: 35. Auf Basis dieser Plasmide wird das Konstrukt pLT12 20 erstellt, indem pUT1 mittels SalI/ScalI geschnitten wird und pUT2 mittels XhoI/ScalI geschnitten wird. Die die Expressionskassetten enthaltenden Fragmente werden ligiert und in E. coli XLI blue MRF transformiert. Es wird nach Vereinzelung Ampicillin-resistenter Kolonien DNA präpariert und per Restriktionsanalyse solche Klone 25 identifiziert, die zwei Expressionskassetten enthalten. Die XhoI/ SalI Ligation kompatibler Enden hat dabei die beiden Schnittstellen XhoI und SalI zwischen den Expressionskassetten eliminiert. Es resultiert das Plasmid pUT12, das in SEQ ID NO: 16 definiert ist. Anschließend wird pUT12 wiederum mittels Sal/ScalI geschnit- 30 ten und pUT3 mittels XhoI/ScalI geschnitten. Die die Expressions- kassetten enthaltenden Fragmente werden ligiert und in E. coli XLI blue MRF transformiert. Es wird nach Vereinzelung Ampicillin- resistenter Kolonien DNA präpariert und per Restriktionsanalyse solche Klone identifiziert, die drei Expressionskassetten enthal- 35 ten. Auf diese Weise wird eine Auswahl von Multiexpressionskas- setten geschaffen, die für die Insertion gewünschter DNA genutzt werden kann und in Tabelle 3 beschrieben wird und zudem noch wei- tere Expressionskassetten aufnehmen kann.

40 Die Expressionskassetten können mehrfach den selben Promotor ent- halten oder aber über drei verschiedene Promotoren aufgebaut wer- den.

Tabelle 4: Multiple Expressionskassetten

	Plasmidname des pUC19-Derivates	Schnittstellen vor dem jeweiligen Promotor	Multiple Klonierungs-Schnittstellen	Schnittstellen hinter dem OCS-Terminator
5	pUT1 (pUC19 mit USP-OCS1)	EcoRI/AscI/SacI/XhoI	(1) BstXI/NotI/PstI/ XbaI/StuI	SalI/EcoRI/SacI/AscI/HindIII
10	pDCT (pUC19 mit DC3-OCS)	EcoRI/AscI/SacI/XhoI	(2) BamHI/EcoRV/ ApaI/NheI/ HpaI	SalI/EcoRI/SacI/AscI/HindIII
15	pLeBT (pUC19-mit LeB4(700)-OCS)	EcoRI/AscI/SacI/XhoI	(3) BglII/NaeI/ ClaI/SmaI/NcoI	SalI/SacI/AscI/HindIII
20	pUDL12 (pUC 19 mit mit USP-OCS1 und mit DC3-OCS)	EcoRI/AscI/SacI/XhoI	(1) BstXI/NotI/ PstI/XbaI/StuI und (2) BamHI/EcoRV/ ApaI/NheI/ HpaI	SalI/EcoRI/SacI/AscI/HindIII
25	pUDL123 Triple expression cassette (pUC19 mit USP/ DC3 und LeB4-700)	EcoRI/AscI/SacI/XhoI	(1) BstXI/NotI/ PstI/XbaI/StuI und (2) BamHI/ (EcoRV*)/ApaI/ NheI/HpaI und (3) BglII/NaeI/ ClaI/SmaI/NcoI	SalI/SacI/AscI/HindIII

* EcoRV Schnittstelle schneidet im 700 Basenpaarfragment des LeB4 Promotors (LeB4-700)

Analog lassen sich weitere Promotoren für Multigenkonstrukte erzeugen insbesondere unter Verwendung des

- a) 2,4 kB Fragmente des LeB4-Promotors (Bäumlein et al., 1991:
30 Mol.Gen.Genet. 225,121-128) oder mithilfe des
- b) Phaseolin-Promotors (Bustos et al. (1989) Plant Cell 1,839-853) oder mithilfe des
- 35 c) konstitutiven v-ATPase c1-Promotors.

Es kann insbesondere wünschenswert sein, weitere besonders geeignete Promotoren zum Aufbau samenspezifischer Multiexpressionskassetten wie z.B. eines der Fragmente des Napin-Promotors (Stalberg et al., 1993: Plant Mol.Biol.23,671-683) oder den Arcelin-5 Promotor (A.Goossens et al., 1999: plant Physiol. 120,1095-1104) zu verwenden.

ii) Erstellung von Expressionskonstrukten in pUC19- oder pGPTV Derivaten, die Promotor und Terminator erhalten und in Kombination mit gewünschten Gensequenzen zur PUFA Gen-expression in pflanzlichen Expressionskassetten enthalten.

5

Multiexpressionskassetten können mittels AscI direkt von pUC19-Derivaten aus Tabelle 3 in den Vektor pGPTV+AscI (siehe iii.)) über die AscI Schnittstelle inseriert werden und stehen zur Inserierung von Zielgenen zur Verfügung. Die entsprechenden 10 Genkonstrukte (pBUT1 ist in SEQUENZ ID NO: 20, pBUT2 ist in SEQUENZ ID NO: 21, pBUT 3 ist in SEQUENZ ID NO: 22, pBUT12 ist in SEQUENZ ID NO: 22 und pBUT123 ist in SEQUENZ ID NO: 24 dargestellt) stehen erfundungsgemäß als Kit zur Verfügung. Alternativ können Gensequenzen in die pUC19 basierten 15 Expressionskassetten inseriert werden und als AscI Fragment in pGPTV+AscI eingesetzt werden.

In pUT12 wird zunächst über BstXI und XbaI die D-6-Elongase Pp_PSE1 in die erste Kassette inseriert. Dann wird die 20 D-6-Desaturase aus Moos (Pp_des6) über BamHI/NaeI in die zweite Kassette inseriert. Es entsteht das Konstrukt pUT-ED. Das AscI Fragment aus dem Plasmid pUT-ED wird in den mit AscI geschnittenen Vektor pGPTV+AscI inseriert und die Orientierung des inserierten Fragmentes mittels Restriktion oder Sequenzierung 25 ermittelt. Es entsteht das Plasmid pB-DHGLA, dessen vollständige Sequenz in SEQUENZ ID NO. 25 dargestellt ist. Die Codierregion der Physcomitrella delta 6 Elongase ist in SEQUENZ ID NO. 26 dargestellt, die der delta 6 Desaturase aus Physcomitrella in SEQUENZ ID NO: 27.

30

In pUT123 wird zunächst über BstXI und XbaI die Δ-6-Elongase Pp_PSE1 in die erste Kassette inseriert. Dann wird die Δ-6-Desaturase aus Moos (Pp_des6) über BamHI/NaeI in die zweite Kassette inseriert und schließlich die Δ-5-Desaturase 35 aus Phaeodactylum (Pt_des5) über BglII in die dritte Kassette inseriert. Das Dreifachkonstrukt erhält den Namen pARA1. Unter Berücksichtigung sequenzspezifischer Restriktionsschnittstellen können weitere Expressionskassetten gemäß Tabelle 5 mit der Bezeichnung pARA2, pARA3 und pARA4 erstellt werden.

40

Das AscI Fragment aus dem Plasmid pARA1 wird in den mit AscI geschnittenen Vektor pGPTV+AscI inseriert und die Orientierung des inserierten Fragmentes mittels Restriktion oder Sequenzierung ermittelt. Die vollständige Sequenz des resultierenden Plasmides 45 pBARA1 ist in SEQUENZ ID NO. 28 dargestellt. Die Codierregion der Physcomitrella delta 6 Elongase ist in SEQUENZ ID NO. 29 dargestellt, die der delta 6 Desaturase aus Physcomitrella

in SEQUENZ ID NO: 30 und die der delta-5 Desaturase aus *Phaeodactylum tricornutum* in SEQUENZ ID NO: 31.

Tabelle 5: Kombinationen von Desaturasen und Elongasen

5

	Gen Plasmid	Δ -6-Desaturase	Δ -5-Desaturase	Δ -6-Elongase
1	PUT-ED	Pp_des6	--	Pp_PSE1
2	pARA1	Pt_des6	Pt_des5	Pp_PSE1
10	pARA2	Pt_des6	Ce_des5	Pp_PSE1
4	pARA3	Pt_des6	Ce_des5	Pp_PSE1
5	pARA4	Ce_des6	Ce_des5	Ce_PSE1
6	PBDHGLA	Pt_des6	--	Pp_PSE1
7	PBARAI	Pt_des6	Pt_des5	Pp_PSE1

15 Plasmide 1 bis 5 sind pUC Derivate, Plasmide 6 bis 7 sind binäre Pflanzentransformationsvektoren

Pp = *Physcomitrella patens*, Pt = *Phaeodactylum tricornutum*
Pp_PSE1 entspricht der Sequenz aus SEQ ID NO: 9.

20 PSE = PUFA spezifische Δ -6-Elongase

Ce_des5 = Δ -5-Desaturase aus *Caenorhabditis elegans* (Genbank Acc.
Nr. AF078796)

Ce_des6 = Δ -6-Desaturase aus *Caenorhabditis elegans elegans*
(Genbank Acc. Nr. AF031477, Basen 11-1342)

25 Ce_PSE1 = Δ -6-Elongase aus *Caenorhabditis elegans* (Genbank Acc.
Nr. AF244356, Basen 1-867)

Auch weitere Desaturasen oder Elongasegensequenzen können in Expressionskassetten beschriebener Art inseriert werden wie
30 z.B. Genbank Acc. Nr. AF231981, NM_013402, AF206662, AF268031,
AF226273, AF110510 oder AF110509.

iii) Transfer von Expressionskassetten in Vektoren zur Transformation von Agrobakterium tumefaciens und zur
35 Transformation von Pflanzen

Chimäre Genkonstrukte auf Basis der in pUC19 beschriebenen können mittels AscI in den binären Vektor pGPTV inseriert. Die multiple Klonierungssequenz wird zu diesem Zweck um eine AscI Schnittstelle erweitert. Zu diesem Zweck wird der Polylinker als zwei doppelsträngige Oligonukleotide neu synthetisiert, wobei eine zusätzliche AscI DNA Sequenz eingefügt wird. Das Oligonukleotid wird mittels EcoRI und HindIII in den Vektor pGPTV inseriert. Es entsteht das Plasmid pGPTV+AscI. Die notwendigen Kloniertechniken sind dem Fachmann bekannt und können einfach wie in Beispiel 1 beschrieben nachgelesen werden.

Beispiel 14: *In vivo*-Mutagenese

Die *in vivo*-Mutagenese von Mikroorganismen kann mittels Passage der Plasmid- (oder einer anderen Vektor-) DNA durch *E. coli* oder andere Mikroorganismen (z.B. *Bacillus spp.* oder Hefen, wie *Saccharomyces cerevisiae*), bei denen die Fähigkeiten, die Unversehrtheit ihrer genetischen Information aufrechtzuerhalten, gestört ist, erfolgen. Übliche Mutator-Stämme haben Mutationen in den Genen für das DNA-Reparatursystem (z.B. *mutHLS*, *mutD*, *mutT* usw.); als Literaturstelle siehe Rupp, W.D. (1996) DNA repair mechanisms, in: *Escherichia coli and Salmonella*, S. 2277-2294, ASM: Washington). Diese Stämme sind dem Fachmann bekannt. Die Verwendung dieser Stämme ist beispielsweise in Greener, A., und Callahan, M. (1994) Strategies 7:32-34, erläutert. Der Transfer mutierter DNA-Moleküle in Pflanzen erfolgt vorzugsweise nach Selektion und Test der Mikroorganismen. Transgene Pflanzen werden nach verschiedenen Beispielen im Beispielteil dieses Dokumentes erzeugt.

20 Beispiel 15: Untersuchung der Expression eines rekombinanten Genproduktes in einem transformierten Organismus

Die Aktivität eines rekombinanten Genproduktes im transformierten Wirtsorganismus kann auf der Transkriptions- und/oder der 25 Translationsebene gemessen werden.

Ein geeignetes Verfahren zur Bestimmung der Menge an Transkription des Gens (ein Hinweis auf die Menge an RNA, die für die Translation des Genproduktes zur Verfügung steht) 30 ist die Durchführung eines Northern-Blots wie unten ausgeführt (als Bezugsstelle siehe Ausubel et al. (1988) Current Protocols in Molecular Biology, Wiley: New York, oder den oben erwähnten Beispielteil), wobei ein Primer, der so gestaltet ist, dass er an das Gen von Interesse bindet, mit einer nachweisbaren 35 Markierung (gewöhnlich radioaktiv oder chemilumineszent) markiert wird, so dass, wenn die Gesamt-RNA einer Kultur des Organismus extrahiert, auf einem Gel aufgetrennt, auf eine stabile Matrix transferiert und mit dieser Sonde inkubiert wird, die Bindung und das Ausmaß der Bindung der Sonde das Vorliegen und auch die Menge 40 der mRNA für dieses Gen anzeigen. Diese Information zeigt den Grad der Transkription des transformierten Gens an. Zelluläre Gesamt-RNA kann aus Zellen, Geweben oder Organen mit mehreren Verfahren, die alle im Fachgebiet bekannt sind, wie zum Beispiel 45 das von Bormann, E.R., et al. (1992) Mol. Microbiol. 6:317-326 beschriebene, präpariert werden.

Northern-Hybridisierung

Für die RNA-Hybridisierung wurden 20 µg Gesamt-RNA oder 1 µg poly(A)⁺-RNA mittels Gelelektrophorese in Agarosegelen mit einer Stärke von 1,25 % unter Verwendung von Formaldehyd, wie beschrieben in Amasino (1986, Anal. Biochem. 152, 304) aufgetrennt, mittels Kapillaranziehung unter Verwendung von 10 x SSC auf positiv geladene Nylonmembranen (Hybond N+, Amersham, Braunschweig) übertragen, mittels UV-Licht immobilisiert und 3 Stunden bei 68°C unter Verwendung von Hybridisierungspuffer (10 % Dextransulfat Gew./Vol., 1 M NaCl, 1 % SDS, 100 mg Herringssperma-DNA) vorhybridisiert. Die Markierung der DNA-Sonde mit dem Highprime DNA labeling-Kit (Roche, Mannheim, Deutschland) erfolgte während der Vorhybridisierung unter Verwendung von alpha-³²P-dCTP (Amersham, Braunschweig, Deutschland). Die Hybridisierung wurde nach Zugabe der markierten DNA-Sonde im gleichen Puffer bei 68°C über Nacht durchgeführt. Die Waschschritte wurden zweimal für 15 min unter Verwendung von 2 X SSC und zweimal für 30 min unter Verwendung von 1 X SSC, 1 % SDS, bei 68°C durchgeführt. Die Exposition der verschlossenen Filter wurde bei -70°C für einen Zeitraum von 1 bis 14 T durchgeführt.

Zur Untersuchung des Vorliegens oder der relativen Menge an von dieser mRNA translatiertem Protein können Standardtechniken, wie ein Western-Blot, eingesetzt werden (siehe beispielsweise Ausubel et al. (1988) Current Protocols in Molecular Biology, Wiley: New York). Bei diesem Verfahren werden die zellulären Gesamt-Proteine extrahiert, mittels Gelelektrophorese aufgetrennt, auf eine Matrix, wie Nitrozellulose, übertragen und mit einer Sonde, wie einem Antikörper, der spezifisch an das gewünschte Protein bindet, inkubiert. Diese Sonde ist gewöhnlich mit einer chemilumineszenten oder kolorimetrischen Markierung versehen, die sich leicht nachweisen lässt. Das Vorliegen und die Menge der beobachteten Markierung zeigt das Vorliegen und die Menge des gewünschten, in der Zelle vorliegenden mutierten Proteins an.

Beispiel 16: Analyse der Auswirkung der rekombinanten Proteine auf die Produktion des gewünschten Produktes

Die Auswirkung der genetischen Modifikation in Pflanzen, Pilzen, Algen, Ciliaten oder auf die Produktion einer gewünschten Verbindung (wie einer Fettsäure) kann bestimmt werden, indem die modifizierten Mikroorganismen oder die modifizierte Pflanze unter geeigneten Bedingungen (wie den vorstehend beschriebenen) gezüchtet werden und das Medium und/oder die zellulären Komponenten auf die erhöhte Produktion des gewünschten Produktes (d.h. von Lipiden oder einer Fettsäure) untersucht wird. Diese Analyse-

101

techniken sind dem Fachmann bekannt und umfassen Spektroskopie, Dünnschichtchromatographie, Färbeverfahren verschiedener Art, enzymatische und mikrobiologische Verfahren sowie analytische Chromatographie, wie Hochleistungs-Flüssigkeitschromatographie

5 (siehe beispielsweise Ullman, Encyclopedia of Industrial Chemistry, Bd. A2, S. 89-90 und S. 443-613, VCH: Weinheim (1985); Fallon, A., et al., (1987) "Applications of HPLC in Biochemistry" in: Laboratory Techniques in Biochemistry and Molecular Biology, Bd. 17; Rehm et al. (1993) Biotechnology, Bd. 3, Kapitel III:

10 "Product recovery and purification", S. 469-714, VCH: Weinheim; Belter, P.A., et al. (1988) Bioseparations: downstream processing for Biotechnology, John Wiley and Sons; Kennedy, J.F., und Cabral, J.M.S. (1992) Recovery processes for biological Materials, John Wiley and Sons; Shaeiwitz, J.A., und Henry, J.D. (1988)

15 Biochemical Separations, in: Ullmann's Encyclopedia of Industrial Chemistry, Bd. B3; Kapitel 11, S. 1-27, VCH: Weinheim; und Dechow, F.J. (1989) Separation and purification techniques in biotechnology, Noyes Publications).

20 Neben den oben erwähnten Verfahren werden Pflanzenlipide aus Pflanzenmaterial wie von Cahoon et al. (1999) Proc. Natl. Acad. Sci. USA 96 (22):12935-12940, und Browse et al. (1986) Analytic Biochemistry 152:141-145, beschrieben extrahiert. Die qualitative und quantitative Lipid- oder Fettsäureanalyse ist beschrieben

25 bei Christie, William W., Advances in Lipid Methodology, Ayr/ Scotland: Oily Press (Oily Press Lipid Library; 2); Christie, William W., Gas Chromatography and Lipids. A Practical Guide - Ayr, Scotland: Oily Press, 1989, Repr. 1992, IX, 307 S. (Oily Press Lipid Library; 1); "Progress in Lipid Research, Oxford:

30 Pergamon Press, 1 (1952) - 16 (1977) u.d.T.: Progress in the Chemistry of Fats and Other Lipids CODEN.

Zusätzlich zur Messung des Endproduktes der Fermentation ist es auch möglich, andere Komponenten der Stoffwechselwege zu analysieren, die zur Produktion der gewünschten Verbindung verwendet werden, wie Zwischen- und Nebenprodukte, um die Gesamteffizienz der Produktion der Verbindung zu bestimmen. Die Analyseverfahren umfassen Messungen der Nährstoffmengen im Medium (z.B. Zucker, Kohlenwasserstoffe, Stickstoffquellen, Phosphat und andere

35 Ionen), Messungen der Biomassezusammensetzung und des Wachstums, Analyse der Produktion üblicher Metabolite von Biosynthesewegen und Messungen von Gasen, die während der Fermentation erzeugt werden. Standardverfahren für diese Messungen sind in Applied Microbial Physiology; A Practical Approach, P.M. Rhodes und P.F.

40 Stanbury, Hrsgb., IRL Press, S. 103-129; 131-163 und 165-192

45 Stanbury, Hrsgb., IRL Press, S. 103-129; 131-163 und 165-192

102

(ISBN: 0199635773) und darin angegebenen Literaturstellen beschrieben.

Ein Beispiel ist die Analyse von Fettsäuren (Abkürzungen: FAME,
5 Fettsäuremethylester; GC-MS, Gas-Flüssigkeitschromatographie-Massenspektrometrie; TAG, Triacylglycerin; TLC, Dünnschicht-chromatographie).

Der unzweideutige Nachweis für das Vorliegen von Fettsäure-
10 produkten kann mittels Analyse rekombinanter Organismen nach Standard-Analyseverfahren erhalten werden: GC, GC-MS oder TLC, wie verschiedentlich beschrieben von Christie und den Literaturstellen darin (1997, in: Advances on Lipid Methodology, Vierter Aufl.: Christie, Oily Press, Dundee, 119-169; 1998, Gaschromatographie-Massenspektrometrie-Verfahren, Lipide 33:343-353).

Das zu analysierende Material kann durch Ultraschallbehandlung, Mahlen in der Glasmühle, flüssigen Stickstoff und Mahlen oder über andere anwendbare Verfahren aufgebrochen werden. Das
20 Material muss nach dem Aufbrechen zentrifugiert werden. Das Sediment wird in Aqua dest. resuspendiert, 10 min bei 100°C erhitzt, auf Eis abgekühlt und erneut zentrifugiert, gefolgt von Extraktion in 0,5 M Schwefelsäure in Methanol mit 2 % Dimethoxypropan für 1 Std. bei 90°C, was zu hydrolysierten Öl-
25 und Lipidverbindungen führt, die transmethylierte Lipide ergeben. Diese Fettsäuremethylester werden in Petrolether extrahiert und schließlich einer GC-Analyse unter Verwendung einer Kapillarsäule (Chrompack, WCOT Fused Silica, CP-Wax-52 CB, 25 mikrom, 0,32 mm) bei einem Temperaturgradienten zwischen 170°C und 240°C für 20 min
30 und 5 min bei 240°C unterworfen. Die Identität der erhaltenen Fettsäuremethylester muss unter Verwendung von Standards, die aus kommerziellen Quellen erhältlich sind (d.h. Sigma), definiert werden.
35 Bei Fettsäuren, für die keine Standards verfügbar sind, muss die Identität über Derivatisierung und anschließende GC-MS-Analyse gezeigt werden. Beispielsweise muss die Lokalisierung von Fettsäuren mit Dreifachbindung über GC-MS nach Derivatisierung mit 4,4-Dimethoxyoxazolin-Derivaten (Christie, 1998, siehe oben)
40 gezeigt werden.

Expressionskonstrukte in heterologen mikrobiellen Systemen

Stämme, Wachstumsbedingungen und Plasmide

5 Der *Escherichia coli*-Stamm XL1 Blue MRF' kan (Stratagene) wurde zur Subklonierung der neuen Desaturase pPDesaturasel aus *Physcomitrella patens* verwendet. Für die funktionelle Expression dieses Gens verwendeten wir den *Saccharomyces cerevisiae*-Stamm INVSc 1 (Invitrogen Co.). *E. coli* wurde in Luria-Bertini-Brühe (LB, 10 Duchefa, Haarlem, Niederlande) bei 37°C kultiviert. Wenn nötig, wurde Ampicillin (100 mg/Liter) zugegeben, und 1,5 % Agar (Gew./Vol.) wurde für feste LB-Medien hinzugefügt. *S. cerevisiae* wurde bei 30°C entweder in YPG-Medium oder in komplettem Minimalmedium ohne Uracil (CMdum; siehe in: Ausubel, F.M., Brent, R., Kingston, 15 R.E., Moore, D.D., Seidman, J.G., Smith, J.A., Struhl, K., Albright, L.B., Coen, D.M., und Varki, A. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York) mit entweder 2 % (Gew./Vol.) Raffinose oder Glucose kultiviert. Für feste Medien wurden 2 % (Gew./Vol.) Bacto™-Agar (Difco) 20 hinzugefügt. Die zur Klonierung und Expression verwendeten Plasmide sind pUC18 (Pharmacia) und pYES2 (Invitrogen Co.).

Beispiel 17: Klonierung und Expression PUFA-spezifischer Desaturasen aus *Phaeodactylum tricornutum*

25 Für die Expression in Hefe wurden die *Phaeodactylum tricornutum*-cDNA-Klone aus Seq ID NO: 1, 3, 5 oder 11 bzw. die Sequenzen aus SEQ ID NO: 7 oder 9 bzw. andere gewünschte Sequenzen zuerst so modifiziert, dass lediglich die Codierregion mittels Polymerase 30 Kettenreaktion unter Zuhilfenahme zweier Oligonukleotide amplifiziert werden. Dabei wurde darauf geachtet, dass eine Konsensussequenz vor dem Startcodon zur effizienten Translation eingehalten wurde. Entweder wurde hierzu die Basenfolge ATA oder AAA gewählt und vor das ATG in die Sequenz eingefügt (Kozak, M. 35 (1986) Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes, Cell 44, 283-292). Vor diesem Konsensustriplett wurde zusätzlich eine Restriktionsschnittstelle eingeführt, die kompatibel sein muss zur Schnittstelle des Zielvektors, 40 in den das Fragment kloniert werden soll und mit dessen Hilfe die Genexpression in Mikroorganismen oder Pflanzen erfolgen soll.

104

Die PCR-Reaktion wurde mit Plasmid-DNA als Matrize in einem Thermocycler (Biometra) mit der Pfu-DNA-(Stratagene)Polymerase und dem folgenden Temperaturprogramm durchgeführt: 3 min bei 96°C, gefolgt von 30 Zyklen mit 30 s bei 96°C, 30 s bei 55°C und 2 min 5 bei 72°C, 1 Zyklus mit 10 min bei 72°C und Stop bei 4°C. Die Anlagerungstemperatur wurde je nach gewählten Oligonukleotiden variiert. Pro Kilobasenpaare DNA ist von einer Synthesezeit von etwa einer Minute auszugehen. Weitere Parameter, die Einfluss auf die PCR haben wie z.B. Mg-Ionen, Salz, DNA Polymerase etc., sind 10 dem Fachmann auf dem Gebiet geläufig und können nach Bedarf variiert werden.

Die korrekte Größe des amplifizierten DNA-Fragments wurde mittels Agarose-TBE-Gelelektrophorese bestätigt. Die amplifizierte DNA 15 wurde aus dem Gel mit dem QIAquick-Gelextraktionskit (QIAGEN) extrahiert und in die SmaI-Restriktionsstelle des dephosphorylierten Vektors pUC18 unter Verwendung des Sure Clone Ligations Kit (Pharmacia) ligiert, wobei die pUC-Derivate erhalten wurden. Nach der Transformation von E. coli XL1 Blue MRF' kan wurde 20 eine DNA-Minipräparation (Riggs, M.G., & McLachlan, A. (1986) A simplified screening procedure for large numbers of plasmid mini-preparation. BioTechniques 4, 310-313) an ampicillinresistenten Transformanden durchgeführt, und positive Klone mittels BamHI-Restriktionsanalyse identifiziert. Die Sequenz des klonierten 25 PCR-Produktes wurde mittels Resequenzierung unter Verwendung des ABI PRISM Big Dye Terminator Cycle Sequencing Ready Reaction Kit (Perkin-Elmer, Weiterstadt) bestätigt.

Δ5 Acyl Lipid desaturase, Pt_des5

30 Primer 1 GAG CTC ACA TAA TGG CTC CGG ATG CGG ATA AGC
Primer 2 CTC GAG TTA CGC CCG TCC GGT CAA GGG

Das PCR-Fragment (1428bp) wurde mithilfe des Sure Clone Kit (Pharmacia) in pUC 18 kloniert, das inserierte Fragment SacI/XhoI 35 verdaut und das Fragment mithilfe entsprechender Restriktions-schnittstellen in pYES2 oder pYES6 inseriert.

Δ6 Acyl Lipid desaturase, Pt_des6

Primer 3 GGA TCC ACA TAA TGG GCA AAG GAG GGG ACG CTC GGG
40 Primer 4 CTC GAG TTA CAT GGC GGG TCC ATC GGG

Das PCR-Fragment (1451 bp) wurde mithilfe des Sure Clone Kit (Pharmacia) in pUC 18 kloniert, das inserierte Fragment BamHI/XhoI verdaut und das Fragment mithilfe entsprechender 45 Restriktionsschnittstellen in pYES2 oder pYES6 inseriert.

105

 Δ 12 Acyl Lipid desaturase, Pt_des12

Primer 5 GGA TCC ACA TAA TGG TTC GCT TTT CAA CAG CC
Primer 6 CTC GAG TTA TTC GCT CGA TAA TTT GC

5 Δ 12 Acyl Lipid desaturase, Pt_des12.2

Primer 7 GGA TCC ACA TAA TGG GTA AGG GAG GTC AAC G
Primer 8 CTC GAG TCA TGC GGC TTT GTT TCG C

Das PCR Fragment (1505bp) wurde mithilfe des Sure Clone

10 Kit (Pharmacia) in pUC 18 kloniert, das inserierte Fragment BamHI/XhoI verdaut und das Fragment mithilfe entsprechender Restriktionsschnittstellen in pYES2 oder pYES6 inseriert.

Die Plasmid-DNA wurde mit Restriktionsenzym/en passend zur
15 eingeführten Schnittstelle der Primersequenz gespalten und das erhaltene Fragment in die kompatiblen Restriktionsstellen des dephosphorylierten Hefe-E. coli-Shuttleektors pYES2 oder pYES6 ligiert, wobei pYES-Derivate erhalten werden. Nach der Transformation von E. coli und DNA-Minipräparation aus den
20 Transformanden wurde die Orientierung des DNA-Fragments im Vektor durch geeignete Restriktionsspaltung oder Sequenzierung überprüft. Ein Klon wurde für die DNA-Maxipräparation mit dem Nucleobond[®] AX 500 Plasmid-DNA-Extraktionskit (Macherey-Nagel, Düringen) angezogen.

25

Saccharomyces cerevisiae INVSc1 wurde mit den pYES-Derivaten und pYES Leervektor mittels eines PEG/Lithiumacetat-Protokolls transformiert (Ausubel et al., 1995). Nach der Selektion auf CMdum-Agarplatten mit 2 % Glucose wurden pYES-Derivate-
30 Transformanden und eine pYES2-Transformande zur weiteren Anzucht und funktionellen Expression ausgewählt. Bei pYES6-Derivaten wurde Blasticidin als Antimetabolit verwendet. Im Fall von Coexpressionen auf Basis von pYES2 und pYES6 wurde auf Minimalmedium mit Blasticidin selektiert.

35

Funktionelle Expression einer Desaturaseaktivität in Hefe

Vorkultur

40 20 ml CMdum-Flüssigmedium ohne Uracil aber mit 2 % (Gew./Vol.) Raffinose wurden mit den transgenen Hefeklonen (pYES2) angeimpft und 3 Tage bei 30°C, 200 rpm gezüchtet, bis eine optische Dichte bei 600 nm (OD_{600}) von 1,5 bis 2 erreicht wurde. Wurde als Vektor pYES6 verwendet, so wurde zusätzlich auf Blasticidin als Anti-
45 metabolit selektiert.

Hauptkultur

Für die Expression wurden 20 ml CMdum-Flüssigmedium ohne Uracil aber mit 2 % Raffinose und 1 % (Vol./Vol.) Tergitol NP-40 5 mit Fettsäuresubstraten auf eine Endkonzentration von 0,003 % (Gew./Vol.) angereichert. Die Medien wurden mit den Vorkulturen auf eine OD₆₀₀ von 0,05 angeimpft. Die Expression wurde bei einer OD₆₀₀ von 0,2 mit 2 % (Gew./Vol.) Galaktose für 16 Std. induziert, wonach die Kulturen eine OD₆₀₀ von 0,8-1,2 geerntet wurden.

10**Fettsäureanalyse**

Die Gesamt-Fettsäuren wurden aus Hefekulturen extrahiert und mittels Gaschromatographie analysiert. Davon wurden Zellen von 15 ml Kultur mittels Zentrifugation (1000 x g, 10 min, 4°C) geerntet und einmal mit 100 mM NaHCO₃, pH 8,0, gewaschen, um restliches Medium und Fettsäuren zu entfernen. Zur Herstellung des Fett-säuremethylester (FAMES oder Singular FAME) wurden die Zell-sedimente mit 1 M methanolischer H₂SO₄ und 2 % (Vol./Vol.) 20 Dimethoxypropan für 1 Std. bei 80°C behandelt. Die FAMES wurden zweimal mit 2 ml Petrolether extrahiert, einmal mit 100 mM NaHCO₃, pH 8,0, und einmal mit destilliertem Wasser gewaschen und mit Na₂SO₄ getrocknet. Das organische Lösungsmittel wurde unter einem Argonstrom verdampft, und die FAMES wurden in 50 mikrol Petrol-ether gelöst. Die Proben wurden auf einer ZEBRON-ZB-Wax-Kapillarsäule (30 m, 0,32 mm, 0,25 mikro m; Phenomenex) in einem Hewlett Packard-6850-Gaschromatograph mit einem Flammenionisationsdetektor aufgetrennt. Die Ofentemperatur wurde von 70°C (1 min halten) bis 200°C mit einer Rate von 20°C/min, dann auf 250°C 30 (5 min halten) mit einer Rate von 5°C/min und schließlich auf 260°C mit einer Rate von 5°C/min programmiert. Stickstoff wurde als Trägergas verwendet (4,5 ml/min bei 70°C). Die Fettsäuren wurden durch Vergleich mit Retentionszeiten von FAME-Standards (SIGMA) identifiziert.

35**Expressionsanalyse**

Die Verhältnisse der zugegebenen und aufgenommenen Fettsäure-substrate wurden ermittelt und so Quantität und Qualität der 40 Desaturasereaktion gemäß Tabelle 6, Tabelle 7 und Tabelle 8 erfasst.

Ergebnis der Expression einer *Phaeodactylum tricornutum* Δ-6-Acyl Lipid Desaturase in Hefe:

45

Tabelle 6

Fettsäure	pYes2	pYes2-Ptd6 gefüttert mit			
		-	+18:2	+18:3	
5	16:0	13,3	18,9	28,4	16,7
	16:1Δ9	45,4	44,7	12,5	16,9
	16:2Δ6,9	-	4,3	-	-
	18:0	4,9	6,3	10,4	9,1
	18:1Δ9	36,4	24,1	6,8	11,8
	18:2Δ6,9	-	1,8	-	-
10	18:2Δ9,12	-	-	33,4	-
	18:3Δ6,12,15	-	-	4,9	-
	18:3Δ9,12,15	-	-		43,1
	18:4Δ6,9,12,15	-	-	-	2,3

15 Die Angaben stellen Mol-% entsprechender cis-Fettsäuren dar.

Ergebnis der Expression einer Phaeodactylum tricornutum Δ-5-Acyl Lipid Desaturase in Hefe:

20 Tabelle 7

Fett-	pYES2	pYES_PtD5-Konstrukt gefüttert mit							20:3	
		Kon-.		Leer	18:2	18:3	20:1	-20:1	20:2	20:3
25	16:0Δ	16,9	20,4	27,7	24,4	16,2	21	17,6	19,5	22,8
	16:1Δ9	44,7	44,1	13,2	9,6	37,4	39,4	38,3	36,9	30,7
	18:0	6,1	6,9	10,54	9,8	4,7	7,9	6,3	6,8	8,2
	18:1Δ9	31,72	28,1	8,77	6	15	26	29,5	25,6	21,1
	18:2Δ5,9	0,17	0	0	0	0,09	0,21	0,09	9	
	18:2Δ9,12	-	39,7	-	-	-	-	-	-	
30	18:3Δ9,12,15	-		49,9	--	-	-	-	-	
	20:1Δ8	-		-	25,5	-	-	-	-	
	20:1Δ11	-		-	-	5,41	-	-	-	
	20:2Δ5,11	-		-	-	-	0,21	-	-	
35	20:2Δ11,14	-	-	-	-	-	-	6,48	-	
	20:3Δ5,11,14	-						0,76	-	
	20:3Δ11,14,17	-	-	-	-	-	-	-	9,83	
	20:3Δ8,11,14	-	-	-	-	-	-	-	-	13,69
40	20:4Δ5,11,14,17	-	-	-	-	-	-	-	1,16	-
	20:4Δ5,8,11,14	-	-	-	-	-	-	-	-	3,08

Die Angaben stellen Mol-% Fettsäuren von cis-Fettsäuren dar.

108

Aus weiteren Fütterungsversuchen wurde gefunden, dass C18:1Δ9 in der Anwesenheit von C18:2 Δ9,11 oder C18:3 Δ9,12,15 oder C20:1 Δ8 Fettsäuren nicht desaturiert wurde während in Anwesenheit von C20:1Δ11, C20:2 Δ11,14 und C20:3 Δ8,11,14 auch C18:1 desaturiert wird. Ebenfalls keine Desaturierung erfolgte in Anwesenheit von C20:3 Δ8,11,14.

Bei Nutzung des Protease-defizienten Hefestammes C13BYS86 (Kunze I. et al., Biochimica et Biophysica Acta (1999) 1410:287-298) für die Expression der Δ-5-Desaturase aus *Phaeodactylum tricornutum* auf Vollmedium mit Blasticidin wurde gefunden, dass C20:4 Δ8,11,14,17 als Substrat der Δ-5-Desaturase mit 20 % Umsatzrate ebenso gut umgesetzt wurde wie C20:3 Δ8,11,14. Alternativ können auch die Auxotrophiemarker leu2, ura3 oder his für Genexpression genutzt werden.

In einem weiteren Coexpressionsexperiment von Δ-5 Desaturase aus *Phaeodactylum* und Δ-6 Elongase aus *Physcomitrella* wurde der Stamm UTL7A (Warnecke et al., J. Biol. Chem. (1999) 274(19):13048-13059) benutzt, wobei die Δ-5 Desaturase ca 10 % C20:3 Δ8,11,14 zu C20:4 Δ5,8,11,14 umsetzte.

Weitere Fütterungsexperimente mit verschiedensten anderen Fettsäuren allein oder in Kombination (z.B. Linolsäure, 20:3 Δ-5,11,14-Fettsäure, alpha- oder gamma Linolensäure, Stearidon-säure, Arachidonsäure, Eicosapentaensäure etc.) können zur detaillierteren Bestätigung der Substratspezifität und -Selektivität dieser Desaturasen durchgeführt werden.

30 Tabelle 8: Ergebnis der Coexpression einer *Phaeodactylum tricornutum* Δ-5-Acyl Lipid Desaturase und einer Δ-6 Elongase aus Moos in Hefe auf Basis der Expressionsvektoren pYes2 und pYes6

	pYes2-Elo		pYes2-Elo and pYes6-Ptd5	
	+18:3	+18:4	+18:3	+18:4
16:0	15,0	14,8	15,6	15,1
16:1Δ9	27,7	29,2	27,5	29,0
18:0	5,6	6,3	5,7	6,4
40 18:1Δ9	17,1	30,8	27,4	31,6
18:3Δ6,9,12	7,60	-	7,8	-
18:4Δ6,9,12,15	-	6,71	-	6,4
20:3Δ8,11,14	15,92	-	13,55	-
45 20:4Δ5,8,11,14	-	-	1,31	-
20:4Δ8,11,14,17	-	11,4	-	10,31
20:5Δ5,8,11,14,17	-	-	-	0,53

109

Aus den Substratumsetzungen geht hervor, dass die verwendete Δ-5-Desaturase aus *Phaeodactylum* und die Δ-6-Elongase aus *Physcomitrella patens* bzgl. der Substrataktivität und insbesondere der Substratspezifität geeignet sind, um Arachidonsäure bzw. Eicosapentaensäure mithilfe erfundungsgemäßer Sequenzen zu produzieren.

Die Fragmentierungsmuster und Massenspektren von DMOX-Derivaten von Standards als auch den Peakfraktionen per GC identifizierter Fettsäuren der in Tabelle 6, 7 und 8 aufgeführten, zeigen vergleichsweise identische Ergebnisse, wodurch die jeweilige Position der Doppelbindung über die bloße GC-Detektion hinaus abgesichert wurde.

15 Beispiel 18: Reinigung des gewünschten Produktes aus transformierten Organismen

Die Gewinnung des gewünschten Produktes aus Pflanzenmaterial oder Pilzen, Algen, Ciliaten, tierischen Zellen oder aus dem Überstand der vorstehend beschriebenen Kulturen kann durch verschiedene, im Fachgebiet bekannte Verfahren erfolgen. Wird das gewünschte Produkt nicht aus den Zellen sezerniert, können die Zellen aus der Kultur durch langsame Zentrifugation geerntet werden, die Zellen können durch Standardtechniken, wie mechanische Kraft oder 25 Ultraschallbehandlung, lysiert werden. Organe von Pflanzen können mechanisch von anderem Gewebe oder anderen Organen getrennt werden. Nach der Homogenisation werden die Zelltrümmer durch Zentrifugation entfernt, und die Überstandsfraktion, welche die löslichen Proteine enthält, wird zur weiteren Reinigung 30 der gewünschten Verbindung aufbewahrt. Wird das Produkt aus gewünschten Zellen sezerniert, werden die Zellen durch langsame Zentrifugation aus der Kultur entfernt, und die Überstandsfraktion wird zur weiteren Reinigung aufbewahrt.

35 Die Überstandsfraktion aus jedem Reinigungsverfahren wird einer Chromatographie mit einem geeigneten Harz unterworfen, wobei das gewünschte Molekül entweder auf dem Chromatographieharz zurückgehalten wird, viele Verunreinigungen in der Probe jedoch nicht, oder die Verunreinigungen auf dem Harz zurückbleiben, die Probe 40 hingegen nicht. Diese Chromatographieschritte können wenn nötig wiederholt werden, wobei die gleichen oder andere Chromatographieharze verwendet werden. Der Fachmann ist in der Auswahl geeigneter Chromatographieharze und ihrer wirksamsten Anwendung für ein bestimmtes zu reinigendes Molekül bewandert. Das 45 gereinigte Produkt kann durch Filtration oder Ultrafiltration konzentriert und bei einer Temperatur aufbewahrt werden, bei der die Stabilität des Produktes maximal ist.

110

Im Fachgebiet ist ein breites Spektrum an Reinigungsverfahren bekannt, und das vorstehende Reinigungsverfahren soll nicht beschränkend sein. Diese Reinigungsverfahren sind zum Beispiel beschrieben in Bailey, J.E., & Ollis, D.F., Biochemical Engineering Fundamentals, McGraw-Hill: New York (1986).

Die Identität und Reinheit der isolierten Verbindungen kann durch Standardtechniken des Fachgebiets bestimmt werden. Dazu gehören Hochleistungs-Flüssigkeitschromatographie (HPLC), spektroskopische Verfahren, Färbeverfahren, Dünnschichtchromatographie, insbesondere Dünnschichtchromatographie und Flammenionisationsdetektion (IATROSCAN, Iatron, Tokio, Japan), NIRS, Enzymtest oder mikrobiologisch. Eine Übersicht über diese Analyseverfahren siehe in: Patek et al. (1994) Appl. Environ. Microbiol. 60:133-140; Malakhova et al. (1996) Biotehnologiya 11:27-32; und Schmidt et al. (1998) Bioprocess Engineer. 19:67-70. Ulmann's Encyclopedia of Industrial Chemistry (1996) Ed. A27, VCH: Weinheim, S. 89-90, S. 521-540, S. 540-547, S. 559-566, 575-581 und S. 581-587; Michal, G (1999) Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology, John Wiley and Sons; Fallon, A., et al. (1987) Applications of HPLC in Biochemistry in: Laboratory Techniques in Biochemistry and Molecular Biology, Bd. 17.

Äquivalente

25

Der Fachmann erkennt oder kann viele Äquivalente der hier beschriebenen erfindungsgemäßen spezifischen Ausführungsformen feststellen, indem er lediglich Routineexperimente verwendet. Diese Äquivalente sollen von den Patentansprüchen umfasst sein.

30

35

40

45

Patentansprüche

1. Expressionskassette mit einer Struktur ausgewählt aus der
5 Gruppe:
 - a) L1 - Promotor - Strukturgen - L2,
 - b) L1 - Promotor - Strukturgen - L2 - L1 - Promotor - Struk-
10 turgen - L2,
 - c) L1 - Promotor - Strukturgen - L2 - L1 - Promotor - Struk-
turgen - L2 - L1 - Promotor - Strukturgen - L2,
- 15 wobei L1, L2, Promotor und Strukturgen die folgende Bedeutung
hat:

L1 = SEQ ID NO: 32 oder eine äquivalente Restriktionsschnitt-
stellen enthaltende Sequenz,
20 L2 = unabhängig voneinander SEQ ID NO: 33, 34 oder 35 oder
äquivalente Restriktionsschnittstellen enthaltende Se-
quenzen,
Promotor = pflanzlicher Promotor
Strukturgen = eine in Pflanzen exprimierbare
25 Nukleinsäuresequenz.
2. Expressionskassette nach Anspruch 1, wobei das Strukturgen
ein Biosynthesegen ist.
- 30 3. Expressionskassette nach Anspruch 1 oder 2, wobei das Struk-
turgen ein Biosynthesegen des Lipid- oder Fettsäurestoffwech-
sels ist.
4. Expressionskassette nach den Ansprüchen 1 bis 3, wobei das
35 Strukturgen ein pflanzliches Gen ist.

40

Zeichn.

45

112

5. Expressionskassette nach den Ansprüche 1 bis 4, wobei das Gen eine Nukleinsäuresequenz ist, die für Proteine ausgewählt aus der Gruppe :
Acyl-CoA-Dehydrogenase(n), Acyl-ACP[= acyl carrier protein]-Desaturase(n), Acyl-ACP-Thioesterase(n), Fettsäure-Acyl-Transferase(n), Fettsäure-Synthase(n), Fettsäure-Hydroxylase(n), Acetyl-Coenzym A-Carboxylase(n), Acyl-Coenzym A-Oxidase(n), Fettsäure-Desaturase(n), Fettsäure-Acetylenasen, Lipoxygenasen, Triacylglycerol-Lipasen, Allenoxid-Synthasen, Hydroperoxid-Lyasen oder Fettsäure-Elongase(n), kodiert.
10. Expressionskassette nach den Ansprüche 1 bis 5, wobei das Gen eine Nukleinsäuresequenz ausgewählt aus der Gruppe ist:
15 a) einer Nukleinsäuresequenz mit der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 11 dargestellten Sequenz,
- 20 b) Nukleinsäuresequenzen, die aufgrund des degenerierten genetischen Codes durch Rückübersetzung der in SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6 oder SEQ ID NO: 12 dargestellten Aminosäuresequenzen erhalten werden,
25 c) Derivate der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 11 dargestellten Nukleinsäuresequenz, die für Polypeptide mit der in SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6 oder SEQ ID NO: 12 dargestellten Aminosäuresequenzen codieren und mindestens 50 % Homologie auf Aminosäureebene aufweisen, ohne daß die enzymatische Wirkung der Polypeptide wesentlich reduziert ist.
- 30 7. Verwendung von Expressionskassetten gemäß den Anspüchen 1 bis 6 in einem Verfahren zur Herstellung von Fettsäureestern mit einem erhöhtem Gehalt an mehrfach ungesättigten Fettsäuren mit mindestens zwei Doppelbindungen dadurch gekennzeichnet, daß man die Expressionskassette in einen Fettsäureester produzierenden Organismus einbringt, anzieht und die in dem Organismus enthaltenen Fettsäureester isoliert.
35 8. Verfahren nach Anspruch 7, wobei die durch das Verfahren hergestellten Fettsäureester mehrfach ungesättigte C₁₈-, C₂₀- oder C₂₂-Fettsäuremoleküle mit mindestens zwei Doppelbindungen im Fettsäureester enthalten.
- 40 45

113

9. Verfahren nach Anspruch 7 oder 8, wobei die C₁₈-, C₂₀- oder C₂₂-Fettsäuremoleküle aus dem Organismus in Form eines Öls, Lipids isoliert werden.
- 5 10. Verfahren nach den Ansprüchen 7 bis 9, wobei der Organismus eine transgene Pflanze ist.
11. Verfahren nach den Ansprüchen 7 bis 10, wobei die Fettsäureester C₁₈-, C₂₀- oder C₂₂-Fettsäuren mit drei, vier oder fünf
10 Doppelbindungen im Fettsäureester enthalten.
12. Verfahren nach den Ansprüchen 7 bis 11, dadurch gekennzeichnet, daß man die in den Fettsäureestern enthaltenden mehrfach ungesättigten Fettsäuren freisetzt.
- 15 13. Vektor enthaltend eine Expressionskassette gemäß den Ansprüchen 1 bis 6.
14. Organismus enthaltend mindestens eine Expressionskassette gemäß den Ansprüchen 1 bis 6 oder mindestens einen Vektor gemäß Anspruch 13.
- 20 15. Organismus nach Anspruch 14, wobei es sich bei dem Organismus um eine Pflanze handelt.
- 25 16. Transgene Pflanze enthaltend eine funktionelle oder nicht funktionelle Nukleinsäuresequenz in einer Expressionskassette gemäß den Ansprüchen 1 bis 6.
- 30 17. Verwendung einer einer Expressionskassette gemäß den Ansprüchen 1 bis 6 zur Herstellung von transgenen Pflanzen.

35

40

45

Figur 1: Polypeptidvergleich der Codierregionen von Pp_des 6 (obere Reihe) mit der EST-Sequenz von PT001078032R (untere Reihe)

```

398 WKPLVWMAVTELMSGMLLGTVFVLSHNGMEVYNSSKEFVSAQI-----VSTR 444
      W+    + + + + L +F LSHN + S+ +A + V T
430 WRVFGNIMLMGVAESLALAVLFLSHN----FESADRDPPTAPLKKTGEPVDFKQTQVETS 263
445 DIKGNIFNDWFTGGGLNRQIEHHLFPTMPRHNLNKIAPRVEVFCKKHGLVY 494
      G + FTGGLN Q+EHHLFPM IAP+V C KHG+ Y
262 CTYGGFLSGCFTGGLNFQVEHHLFPRMSSAWYXYIAPKVREICAKHGVHY 113

```

Figur 2: Polypeptidvergleich von Codierregionen der Ma_des12 (obere Sequenz) mit PT001070010R

```

105 GVVVLAHECGHQSFSTSCTLNN 126
      G WVLAHECGH +FS +++L +
533 GFWVLAHECGHGAFSKNRSLQD 598

```

Figur 2a: Polypeptidvergleich von Codierregionen der Ma_des12 (obere Sequenz) mit PT001072031R

```

117 SFSTSCTLNNTVGWILHSMLLPYHWSRISHSKHH 151
      ++S S+T N+ VG+I+H LLVPY +W+ +H+KHH
465 AYSDSQTFNDVVGFIHVQALLVPYFAWQYTHAKHH 569

```

Figur 3: Polypeptidvergleich von Codierregionen eines PCR Produktes aus Primerpaar F6a und R4a2 codierend für ein Desaturase Fragment (obere Reihe) aus Phaodactylum mit der Sequenz T36617 aus Streptomyces coelicolor (untere Reihe)

```

1   WWKNKHNGHHAVPNLHCSSAVAQDGDPIDTMPLLAWSVQQAQSYRELQADGKDGLVKF 60
      WW++KH H-HA PN +D DPDI LL WS QA++ +GL +
114 WWQDKHTRHHANPN-----TEDLDPDIGP-DLLVWSPDQARAA-----TGLPRL 156
61 MIRNQSYFYFPILLARLSWLNEFKCAFGLGAASENAALELKAKGLQYPLLEKAGILLH 120
      + R Q++ +FP+L L E F G A N L+ +A L+ A +L H
157 LGRWQAFLFFPLTL-----EGFNLHVASGRAMANRRLKRRA-----LDGALLLAH 202
121 YAWMLTVSSGFGRXXXXXXXXXXXXXXCGFLLAIVFGLGHNGMATYNADARPDFWKLQ 180
      A LT F G L F H GM AD RPDF + Q
203 CAVYLTA--FWVLPPGMAIAFLAVHQCLFGVYLGSAFAPNHKGMPILTADDRPDFLRRQ 260
181 VTTTRNVTGGHGFQAFVDWFCGGLQYQVDHHLFPS 216
      V T+RNV GG F D GGL +Q++HHLFPS
261 VLTSRNVNGG-----LFTDLALGGLNHQIEHHLFPS 291

```

Figur 4: Polypeptidvergleich von Codierregionen aus Pp_des6 (obere Reihe) verglichen mit Pt_des6 (untere Reihe)

51 KRLTSKKRVSESAAVQCISAEVQRNSSTQGTAEALAESVVKPTRRRSSQW 100
.....|.....|.....
1MGKGGDARASKG 12

101 KKSTHPLS..EVAVHNKPSCDWIVVKNKVYDVSNFADEHPGGSVISTYFG 148
.....:||| | | |||:|||||||.|||||.|||:
13 STAARKISWQEVKTHASPEDAWIIHSNKVYDVSNW.HEHPGGAVIFTHAG 61

149 RDGTDVFSSFHAASTWKILQDFYIGDV..ERVEPTPELL...KDFREMRA 193
| |||:..||| . .::| |||:. | | | :| :| :|||:
62 DDMTIDIFAAPHAPGSQSLMKKFYIGELLPETTGKEPQQIAFEKGYRDLRS 111

194 LFLREQLFKSSKLYYVMKLLTNVAIFAASIAIICWSKTISAVLASACMMA 243
: : |||.| :|| | | .|||.|||.||:: :| | ||| |:
112 KLIMMGFKSNKWFYVYKCLSNSMIAWAAACALVFYSDRFWVHLASAVMLG 161

244 LCFQQCGWLSHDFLHNQVFETRWLNEVVGYVIGNAVLGFSTGWWKEKHNL 293
||| |||..|||||.||| | : .| | ||| .|:| | ||| |||
162 TFFQQSGWLAHDFLHHQVFTKRKHGDLGLFWGNLMOQGYSVQWNKHN 211

294 HHAAPN.ECDQTY.QPIDEDIDTLPLIAWS.....KDILATVENKTFL 334
||| ||| | | | | |||:|||:||| ::| .. .
212 HHAVPNLHCSSAVAQDGDPIDTMPLLAWSVQQAQSYYRELQADGKDSGLV 261

335 R.IIQLYQHLFFMGLFFARGSWLFWSWR.....YTSTAVLSPVDR... 373
: .: | | : :| | ||| | .: . . | | | | :
262 KFMIRNQSYFYFPILLARLSWLNESFKCAFGLGAASENAALELKAKGLQ 311

374 ..LLEKGTVLFHYFWFVGTCAC.YLLPGWKPLVWMAVTTELMS.GMLLGFVF 419
| ||| :| ||| | . . : . . . | | | | ||| |||
312 YPLLEKAGILLHYAWMLTVSSGFGRFSFAYTAFYFLTATASCGFLLAIVF 361

420 VLSHNGMEVNSS..KEFVSAQIVSTRDIKG....NIFNDWFTGGLNRQ 462
| | ||| | |. :| | : .|||. | | | | ||| |||
362 GLGHNGMATYNADARPFWKLQVTTTRNVTGGHGFQAFVDWFCGGLQYQ 411

463 IEHHLFPTMPRHNLNKIAPRVEVFCKKHGLVYEDVSIATGTCVKLALK 512
::| ||| | .:| ||| | | | |||. | . | : : | | .|||
412 VDHHLFPSLPRHNLAKTHALVESFCKEWGVQYHEADLVDGTMELHHLGS 461

513 VAEAAAEEQHATTS.... 525
|| |
462 VAGFFFVVDFVRDGPAM. 477

Figur 5: Polypeptidvergleich von Codierregionen aus Pp_des6 (obere Reihe) verglichen mit Pt_des5 (untere Reihe)

```

51 KRLTSKKRVSESAAVQCISAEVQRNSSTQGTAEALAESVVKPTRRRSSQW 100
      :|    | .|...
1 .. .... MAPDADKLRQRQTTAV 16

101 KKSTHPLSEVAVHNKPSDC.....WIVVKNKVYDVSNFADEHPGGSVIS 144
     | | . : : : : || . | :||| | |
17 AK..HNAATISTQERLCSLSSLKGEEVCIDGIIYDLQSF..DHPGGETIK 62

145 TYFGRDGTDVFSSFHAASTWKILQDF.YIGDVERVEPTPELLKDF.REM. 191
     : | | | : | | | : | : | | . : | | |
63 MFGGNDVTQYKMIHPYHTEKHLEKMKRUVGKVTDVCEYKFDTEFEREIK 112

192 RALFLREQLFKS.SKLYYVMKLLTNVAIFAAASIAIIICWSKT.ISAVLASA 239
     | . | . | : : : || | | | | . || |
113 REVFKIVRRGKDFGTGLWFFRAFCYIAIF..FYLQYHWVTTGTSWLLAVA 160

240 CMMALCFQQCGWLSHDFLHNQVFETRWLNEVVGVYVIGNAVLGFSTGWWKE 289
     .. . | | . | . | . | : | : | | | | . |
161 YGISQAMIGMN.VQHDANHGATSKRPWVNDMLG..LGADFIGGSKWLIQE 207

290 KHNLHHAAPNECDQTYQPIDEDEDITLPLIAWSKDILATVENKTFRLILQY 339
     . | | | | : | : | : | : | : | . | .. |
208 QHWTHHAYTNHAEM..DP..DSFGAEPMLLFN.DYPLDHPARTWLH..RF 250

340 QHLFFMGLLFFARGSWLFWWR.....YTSTAVLS...PVDRLLEKGTVL 381
     | | : | | | | | . | | | . | | | . |
251 QAFFYMPVL...AGYWLSAVFPNPQILDLQQRGALSGVIRLDNAFIHSRRK 297

382 FHYFW...FVG....TACYLLPG....WKPLUWMMAVTELMSGMILLGFVFV 420
     : | | : | | | | : | . . . : | . |
298 YAVFWRAVYIAVNVIAPFYTNGLEWSWRVFGNIMLMGVAESLALAVLFS 347

421 LSHN.....GMEVYNSSKEFVSAQIVSTRDIKGNIIFNDWFTGGLN 460
     | | | | : | . | | | | | . | | | | |
348 LSHNFBSADRDPPTAPLKKTGEPVDFWFKTQ.VETSCTYGGFLSGCFTGGLN 396

461 RQIEHHLFPTMPRHNLNKIAPRVEVFCKKHGLVYEDVS.IATGTCKVLKA 509
     | : | | | | | | | : | | | | | | | . |
397 FQVEHHLFPRMSSAWPYIAPKREICAKHGVHYAYYPWIHQNFLSTVRY 446

510 LKEVAEAAA.EQHATTS..... 525
     : | | | . | |
447 MHAAGTGANWRQMARENPLTGRA. 469

```

Figur 6: Polypeptidvergleich von Codierregionen der Δ -12-Desaturase aus *Mortierella alpina* (Ma_des12) obere Reihe mit der homologen Sequenz aus *Phaeodactylum tricornutum* (Pt_des12) in der unteren Reihe

40 KEIRECIPAHCFERSGLRGLCHVAIDLTLWASLL..FLAATQIDKF..NP 85
|::| || ||| : | ... | .:| . | : : ||
107 KDLRAVIPKDCFEPDTAKSLGYLESVS.TMGTILCSVVGANLLSVLDPSNP 155
.
.
86 LIRYLAWPVYWMQGIVCTGVWVLAHECGHQSFSTSCTLNNNTVGWILHSM 135
| : | | . | | . |||||||| . || ..| . ||:|:
156 L.TWPLWAAYGAVTGTVAMGLWVLAHECGHGAFSKNRSLQDAVGYIHSI 204
.
.
136 LLV?YHSWRISHSKHHKATGHMTKDQVFVPKTRSQVGLPPKENAAAQVE 185
:||||||.||.||.||| : || . | | | | | |
205 MLVPYFSWQRSHAVHHQYTINHMELETHVPDRADKEG....EKSLALRQF 250
.
.
186 EDMSVHLDEEAPIVTLFWMVIQFLFGWPAYLIMNASGQDYGRWTSHFHTY 235
| | . : : |||||. | | | .||:
251 MLDSFGKDKGMKAYGGLQSFLHLIVGWPAYLLIGATGGPDRGMTNHFPYP. 299
.
.
236 SPIFEP....RNFF.....DIIISDLGVLAALGALIYASMQQLSLLTVTK 275
.||: | : | : | : ||:|: | .||| . | | |
300 NPLSTPTQPKKELFPGNWKEKVYQSDIGIAAVVGALIAWTATSGLAPVMA 349
.
.
276 YYIVPYLFVNFWLVLITFLQHTDPKLPHYREGAWNFRQGALCTVDRSGFK 325
| | : : | |||| .||| .||: | | : || | : ||
350 LYGGPLIVINAWLVLYTWLQHTTDVPHFSSDNHNFVKGALHTIDRPYDK 399
.
.
326FLDHMFHGIVHTHVAHHLFQSQMPFYHAEEATYHLKKLLGEYYVD 370
:| : | | ||||| | . | | : | | : | | .|||
400 LDPWGIIDFLHHKIGTTHVAHHFDSTIPHYKAQIATDAIKAKFPEVYLYD 449
.
.
371 PSPIVVAVWRSFRECRFVEDQGDVVFFK 398
| .|| | .|| : | | | .|| . |
450 PTPIPQAMWRVAKGCTAVEQRGDAWWK 477

Figur 7: Polypeptidvergleich von Codierregionen der Δ -12-Desaturase aus *Mortierella alpina* (Ma_des12) obere Reihe mit der homologen Sequenz aus *Phaeodactylum tricornutum* Klon PT001072031R (Pt_des12.2) in der unteren Reihe.

22 NSAKPAFERNYQLPEFTIKEIRECIPAHCFERSGLRGLCHVAIDLWTWASL 71
..| | . . . :|| | :|| || || :|| .. || | .
33 SSYNPLAKDSPELP..TKGQIKAVIPKECFQRSAFWSTFYLMRDILAMAAA 80

72 LFLAATQIDKFENP.....LIRYLAWPVYWIMQGIVCTGVWVLAHECGH 115
.|| :| | | | || || :|| || .|||||
81 FCYGTSQLSTDLPQDATLILPWALGWGVYAFWMGTILTGPWVVAHECGH 130

116 QSFSTSCKTLNNNTVGWILHSMLLVYHSWRISHSKHHKATGHMTKDQVFVP 165
.|| | .|| .|| .|| | || | .|| .|| || :|| :|| :||
131 GAYSDSQTFNDVVGFIVHQALLVPYFANQYTHAKHHRRTNHLVDGESHVP 180

166 KTRSQVGLPP..KENAAAAVQEEDMSVHLDEEAPIVTLFWMVIQFLFGWP 213
| | || | . . | | | | :| | | | || |
181 STAKDNGLGPHERNSFYAAWHEAMG....DGAFAVFQVWS..HLFVGWP 224

214 AYLI.MNASGQ..DYGRW.....TSHFHTYSPIFEPRNFFDIIISDLG 253
|| :|| . | | | | || || .|| :| :|| :||
225 LYLAGLASTGKLAHEGWLEERNAIADHFRPSSPMFFPAKIRAKIALSSAT 274

254 VLAALGALIYASMQSLLTVTKYIVPYLFVNFWLVLITFLQHTDPKLPH 303.
|| | | :|| | . | | :|| || || || || | .||||| :||
275 ELAVLAGLLYVGTQVGHPVLLWYGPYTFVNAWLVLYTTLQHTDPSIPH 324

304 YREGAWNFORGALCTVDRSGFKFLDHMFHGIVHTHVAHHLFQSQMPFYHAE 353
| | | | . :|| | :|| :| | | | | | || || || || :|| .||
325 YGEGETWVKGALSTIDRDY GIF.DFFHHTIGSTHVHHLFHEMPWYNAG 373

354 EATYHLKKLLGE..YYVYDPSPIVVAVWRSFRECRFVEDQGDVFFK 398
|| .|. | | || .| .|| | | :|| | :||
374 IATOKVKEFLEPOGLNYDPTPWyKAMWRIARTCHYVESNEGVQYFK 420

SEQUENZPROTOKOLL

<110> BASF Plant Science GmbH

<120> Verfahren zur Expression von Biosynthesegenen in pflanzlichen Samen unter Verwendung von neuen multiplen Expressionskonstrukten

<130> 2000_904

<140> 2000_904

<141> 2000-12-22

<160> 35

<170> PatentIn Vers. 2.0

<210> 1

<211> 1652

<212> DNA

<213> Phaeodactylum tricornutum.

<220>

<221> CDS

<222> (115) ..(1524)

<400> 1

gacccaacaa acccaacaat cccaaacaaatccatcaaacag gaattgggt tcgttgcgtc 60

aataattgct agaatccaaa cagacagaca gagaccaacc gcattttat caga atg 117
Met
1gct ccg gat gat aag ctt cga caa cgc cag acg act gct gta gct 165
Ala Pro Asp Ala Asp Lys Leu Arg Gln Arg Gin Thr Thr Ala Val Ala
5 10 15aag cac aat gct acc ata tcg acg cag gaa cgc ctt tgc agt ctg 213
Lys His Asn Ala Ala Thr Ile Ser Thr Gln Glu Arg Leu Cys Ser Leu
20 25 30tct tcg ctc aaa ggc gaa gaa gtc tgc atc gac gga atc atc tat gac 261
Ser Ser Leu Lys Gly Glu Val Cys Ile Asp Gly Ile Ile Tyr Asp
35 40 45ctc caa tca ttc gat cat ccc ggg ggt gaa acg atc aaa atg ttt ggt 309
Leu Gln Ser Phe Asp His Pro Gly Gly Glu Thr Ile Lys Met Phe Gly
50 55 60 65ggc aac gat gtc act gta cag tac aag atg att cac ccg tac cat acc 357
Gly Asn Asp Val Thr Val Gln Tyr Lys Met Ile His Pro Tyr His Thr
70 75 80gag aag cat ttg gaa aag atg aag cgt gtc ggc aag gtg acg gat ttc 405
Glu Lys His Leu Glu Lys Met Lys Arg Val Gly Lys Val Thr Asp Phe
85 90 95

gtc tgc gag tac aag ttc gat acc gaa ttt gaa cgc gaa atc aaa cga 453

Val Cys Glu Tyr Lys Phe Asp Thr Glu Phe Glu Arg Glu Ile Lys Arg			
100	105	110	
gaa gtc ttc aag att gtg cga cga ggc aag gat ttc ggt act ttg gga			501
Glu Val Phe Lys Ile Val Arg Arg Gly Lys Asp Phe Gly Thr Leu Gly			
115	120	125	
tgg ttc ttc cgt gcg ttt tgc tac att gcc att ttc ttc tac ctg cag			549
Trp Phe Phe Arg Ala Phe Cys Tyr Ile Ala Ile Phe Phe Tyr Leu Gln			
130	135	140	145
tac cat tgg gtc acc acc gga acc tct tgg ctg ctg gcc gtc gac tac			597
Tyr His Trp Val Thr Thr Gly Thr Ser Trp Leu Leu Ala Val Ala Tyr			
150	155	160	
gga atc tcc caa gcg atg att ggc atg aat gtc cag cac gat gcc aac			645
Gly Ile Ser Gln Ala Met Ile Gly Met Asn Val Gln His Asp Ala Asn			
165	170	175	
cac ggg gcc acc tcc aag cgt ccc tgg gtc aac gac atg cta ggc ctc			693
His Gly Ala Thr Ser Lys Arg Pro Trp Val Asn Asp Met Leu Gly Leu			
180	185	190	
ggt gcg gat ttt att ggt ggt tcc aag tgg ctc tgg cag gaa caa cac			741
Gly Ala Asp Phe Ile Gly Gly Ser Lys Trp Leu Trp Gln Glu Gln His			
195	200	205	
tgg acc cac cac gct tac acc aat cac gcc gag atg gat ccc gat agc			789
Trp Thr His His Ala Tyr Thr Asn His Ala Glu Met Asp Pro Asp Ser			
210	215	220	225
ttt ggt gcc gaa cca atg ctc cta ttc aac gac tat ccc ttg gat cat			837
Phe Gly Ala Glu Pro Met Leu Leu Phe Asn Asp Tyr Pro Leu Asp His			
230	235	240	
ccc gct cgt acc tgg cta cat cgc ttt caa gca ttc ttt tac atg ccc			885
Pro Ala Arg Thr Trp Leu His Arg Phe Gln Ala Phe Phe Tyr Met Pro			
245	250	255	
gtc ttg gct gga tac tgg ttg tcc gct gtc ttc aat cca caa att ctt			933
Val Leu Ala Gly Tyr Trp Leu Ser Ala Val Phe Asn Pro Gln Ile Leu			
260	265	270	
gac ctc cag caa cgc ggc gca ctt tcc gtc ggt atc cgt ctc gac aac			981
Asp Leu Gln Gln Arg Gly Ala Leu Ser Val Gly Ile Arg Leu Asp Asn			
275	280	285	
gct ttc att cac tcg cga cgc aag tat gcg gtt ttc tgg cgg gct gtg			1029
Ala Phe Ile His Ser Arg Arg Lys Tyr Ala Val Phe Trp Arg Ala Val			
290	295	300	305
tac att gcg gtg aac gtg att gct ccg ttt tac acä aac tcc ggc ctc			1077
Tyr Ile Ala Val Asn Val Ile Ala Pro Phe Tyr Thr Asn Ser Gly Leu			
310	315	320	
gaa tgg tcc tgg cgt gtc ttt gga aac atc atg ctc atg ggt gtg gcg			1125
Glu Trp Ser Trp Arg Val Phe Gly Asn Ile Met Leu Met Gly Val Ala			
325	330	335	

gaa tcg ctc gcg ctg gtc ctg ttt tcg ttg tcg cac aat ttc gaa 1173
 Glu Ser Leu Ala Leu Ala Val Leu Phe Ser Leu Ser His Asn Phe Glu
 340 345 350

tcc gcg gat cgc gat ccg acc gcc cca ctg aaa aag acg gga gaa cca 1221
 Ser Ala Asp Arg Asp Pro Thr Ala Pro Leu Lys Lys Thr Gly Glu Pro
 355 360 365

gtc gac tgg ttc aag aca cag gtc gaa act tcc tgc act tac ggt gga 1269
 Val Asp Trp Phe Lys Thr Gln Val Glu Thr Ser Cys Thr Tyr Gly Gly
 370 375 380 385

ttc ctt tcc ggt tgc ttc acg gga ggt ctc aac ttt cag gtt gaa cac 1317
 Phe Leu Ser Gly Cys Phe Thr Gly Gly Leu Asn Phe Gln Val Glu His
 390 395 400

cac ttg ttc cca cgc atg agc agc gct tgg tat ccc tac att gcc ccc 1365
 His Leu Phe Pro Arg Met Ser Ser Ala Trp Tyr Pro Tyr Ile Ala Pro
 405 410 415

aag gtc cgc gaa att tgc gcc aaa cac ggc gtc cac tac gcc tac tac 1413
 Lys Val Arg Glu Ile Cys Ala Lys His Gly Val His Tyr Ala Tyr Tyr
 420 425 430

ccg tgg atc cac caa aac ttt ctc tcc acc gtc cgc tac atg cac gcg 1461
 Pro Trp Ile His Gln Asn Phe Leu Ser Thr Val Arg Tyr Met His Ala
 435 440 445

gcc ggg acc ggt gcc aac tgg cgc cag atg gcc aga gaa aat ccc ttg 1509
 Ala Gly Thr Gly Ala Asn Trp Arg Gln Met Ala Arg Glu Asn Pro Leu
 450 455 460 465

acc gga cgg gcg taa aagtacacga cacgaccaaa ggtggcgtat ggtgatctct 1564
 Thr Gly Arg Ala
 470

agaaaacaga catagcctac tggaaatatac gacgtccaaa caataattt aaagactatt 1624

tttctgcgtta aaaaaaaaaa aaaaaaaa 1652

<210> 2
 <211> 469
 <212> PRT
 <213> Phaeodactylum tricornutum

<400> 2
 Met Ala Pro Asp Ala Asp Lys Leu Arg Gln Arg Gln Thr Thr Ala Val
 1 5 10 15

Ala Lys His Asn Ala Ala Thr Ile Ser Thr Gln Glu Arg Leu Cys Ser
 20 25 30

Leu Ser Ser Leu Lys Gly Glu Glu Val Cys Ile Asp Gly Ile Ile Tyr
 35 40 45

Asp Leu Gln Ser Phe Asp His Pro Gly Glu Thr Ile Lys Met Phe

50	55	60
Gly Gly Asn Asp Val Thr Val Gln Tyr Lys Met Ile His Pro Tyr His		
65	70	75
Thr Glu Lys His Leu Glu Lys Met Lys Arg Val Gly Lys Val Thr Asp		
85	90	95
Phe Val Cys Glu Tyr Lys Phe Asp Thr Glu Phe Glu Arg Glu Ile Lys		
100	105	110
Arg Glu Val Phe Lys Ile Val Arg Arg Gly Lys Asp Phe Gly Thr Leu		
115	120	125
Gly Trp Phe Phe Arg Ala Phe Cys Tyr Ile Ala Ile Phe Phe Tyr Leu		
130	135	140
Gln Tyr His Trp Val Thr Thr Gly Thr Ser Trp Leu Leu Ala Val Ala		
145	150	155
Tyr Gly Ile Ser Gln Ala Met Ile Gly Met Asn Val Gln His Asp Ala		
165	170	175
Asn His Gly Ala Thr Ser Lys Arg Pro Trp Val Asn Asp Met Leu Gly		
180	185	190
Leu Gly Ala Asp Phe Ile Gly Gly Ser Lys Trp Leu Trp Gln Glu Gln		
195	200	205
His Trp Thr His His Ala Tyr Thr Asn His Ala Glu Met Asp Pro Asp		
210	215	220
Ser Phe Gly Ala Glu Pro Met Leu Leu Phe Asn Asp Tyr Pro Leu Asp		
225	230	235
240		
His Pro Ala Arg Thr Trp Leu His Arg Phe Gln Ala Phe Phe Tyr Met		
245	250	255
Pro Val Leu Ala Gly Tyr Trp Leu Ser Ala Val Phe Asn Pro Gln Ile		
260	265	270
Leu Asp Leu Gln Gln Arg Gly Ala Leu Ser Val Gly Ile Arg Leu Asp		
275	280	285
Asn Ala Phe Ile His Ser Arg Arg Lys Tyr Ala Val Phe Trp Arg Ala		
290	295	300
Val Tyr Ile Ala Val Asn Val Ile Ala Pro Phe Tyr Thr Asn Ser Gly		
305	310	315
320		
Leu Glu Trp Ser Trp Arg Val Phe Gly Asn Ile Met Leu Met Gly Val		
325	330	335
Ala Glu Ser Leu Ala Leu Ala Val Leu Phe Ser Leu Ser His Asn Phe		
340	345	350
Glu Ser Ala Asp Arg Asp Pro Thr Ala Pro Leu Lys Lys Thr Gly Glu		
355	360	365

Pro Val Asp Trp Phe Lys Thr Gln Val Glu Thr Ser Cys Thr Tyr Gly
 370 375 380

Gly Phe Leu Ser Gly Cys Phe Thr Gly Gly Leu Asn Phe Gln Val Glu
 385 390 395 400

His His Leu Phe Pro Arg Met Ser Ser Ala Trp Tyr Pro Tyr Ile Ala
 405 410 415

Pro Lys Val Arg Glu Ile Cys Ala Lys His Gly Val His Tyr Ala Tyr
 420 425 430

Tyr Pro Trp Ile His Gln Asn Phe Leu Ser Thr Val Arg Tyr Met His
 435 440 445

Ala Ala Gly Thr Gly Ala Asn Trp Arg Gln Met Ala Arg Glu Asn Pro
 450 455 460

Leu Thr Gly Arg Ala
 465

<210> 3

<211> 1434

<212> DNA

<213> Phaeodactylum tricornutum

<220>

<221> CDS

<222> (1)..(1434)

<400> 3

atg ggc aaa gga ggg gac gct cgg gcc tcg aag ggc tca acg gcg gct 48
 Met Gly Lys Gly Gly Asp Ala Arg Ala Ser Lys Gly Ser Thr Ala Ala
 1 5 10 15

cgc aag atc agt tgg cag gaa gtc aag acc cac gcg tct ccg gag gac 96
 Arg Lys Ile Ser Trp Gln Glu Val Lys Thr His Ala Ser Pro Glu Asp
 20 25 30

gcc tgg atc att cac tcc aat aag gtc tac gac gtg tcc aac tgg cac 144
 Ala Trp Ile Ile His Ser Asn Lys Val Tyr Asp Val Ser Asn Trp His
 35 40 45

gaa cat ccc gga ggc gcc gtc att ttc acg cac gcc ggt gac gac atg 192
 Glu His Pro Gly Gly Ala Val Ile Phe Thr His Ala Gly Asp Asp Met
 50 55 60

acg gac att ttc gct gcc ttt cac gca ccc gga tcg cag tcg ctc atg 240
 Thr Asp Ile Phe Ala Ala Phe His Ala Pro Gly Ser Gln Ser Leu Met
 65 70 75 80

aag aag ttc tac att ggc gaa ttg ctc ccg gaa acc acc ggc aag gag 288
 Lys Lys Phe Tyr Ile Gly Glu Leu Leu Pro Glu Thr Thr Gly Lys Glu
 85 90 95

ccg cag caa atc gcc ttt gaa aag ggc tac cgc gat ctg cgc tcc aaa 336

Pro Gln Gln Ile Ala Phe Glu Lys Gly Tyr Arg Asp Leu Arg Ser Lys			
100	105	110	
ctc atc atg atg ggc atg ttc aag tcc aac aag tgg ttc tac gtc tac			384
Leu Ile Met Met Gly Met Phe Lys Ser Asn Lys Trp Phe Tyr Val Tyr			
115	120	125	
aag tgc ctc agc aac atg gcc att tgg gcc gcc gac tgc ttt gtc gtc			432
Lys Cys Leu Ser Asn Met Ala Ile Trp Ala Ala Ala Cys Ala Leu Val			
130	135	140	
ttt tac tcg gac cgc ttc tgg gta cac ctg gcc agc gcc gtc atg ctg			480
Phe Tyr Ser Asp Arg Phe Trp Val His Leu Ala Ser Ala Val Met Leu			
145	150	155	160
gga aca ttc ttt cag cag tcg gga tgg ttg gca cac gac ttt ctg cac			528
Gly Thr Phe Phe Gln Gln Ser Gly Trp Leu Ala His Asp Phe Leu His			
165	170	175	
cac cag gtc ttc acc aag cgc aag cac ggg gat ctc gga gga ctc ttt			576
His Gln Val Phe Thr Lys Arg Lys His Gly Asp Leu Gly Gly Leu Phe			
180	185	190	
tgg ggg aac ctc atg cag ggt tac tcc gta cag tgg tgg aaa aac aag			624
Trp Gly Asn Leu Met Gln Gly Tyr Ser Val Gln Trp Trp Lys Asn Lys			
195	200	205	
cac aac gga cac cac gcc ccc aac ctc cac tgc tcc tcc gca gtc			672
His Asn Gly His His Ala Val Pro Asn Leu His Cys Ser Ser Ala Val			
210	215	220	
gcg caa gat ggg gac ccg gac atc gat acc atg ccc att ctc gcc tgg			720
Ala Gln Asp Gly Asp Pro Asp Ile Asp Thr Met Pro Leu Leu Ala Trp			
225	230	235	240
tcc gtc cag caa gcc cag tct tac ccg gaa ctc caa gcc gac gga aag			768
Ser Val Gln Gln Ala Gln Ser Tyr Arg Glu Leu Gln Ala Asp Gly Lys			
245	250	255	
gat tcg ggt ttg gtc aag ttc atg atc cgt aac caa tcc tac ttt tac			816
Asp Ser Gly Leu Val Lys Phe Met Ile Arg Asn Gln Ser Tyr Phe Tyr			
260	265	270	
ttt ccc atc ttg ttg ctc gcc cgc ctg tcg tgg ttg aac gag tcc ttc			864
Phe Pro Ile Leu Leu Ala Arg Leu Ser Trp Leu Asn Glu Ser Phe			
275	280	285	
aag tgc gcc ttt ggg ctt gga gct ggc tcg gag aac gct gct ctc gaa			912
Lys Cys Ala Phe Gly Leu Gly Ala Ala Ser Glu Asn Ala Ala Leu Glu			
290	295	300	
ctc aag gcc aag ggt ctt cag tac ccc ctt ttg gaa aag gct ggc atc			960
Leu Lys Ala Lys Gly Leu Gln Tyr Pro Leu Leu Glu Lys Ala Gly Ile			
305	310	315	320
ctg ctg cac tac gct tgg atg ctt aca gtt tcg tcc ggc ttt gga cgc			1008
Leu Leu His Tyr Ala Trp Met Leu Thr Val Ser Ser Gly Phe Gly Arg			
325	330	335	

ttc tcg ttc gcg tac acc gca ttt tac ttt cta acc gcg acc gcg tcc Phe Ser Phe Ala Tyr Thr Ala Phe Tyr Phe Leu Thr Ala Thr Ala Ser 340	345	350	1056	
tgt gga ttc ttg ctc gcc att gtc ttt ggc ctc ggc cac aac ggc atg Cys Gly Phe Leu Leu Ala Ile Val Phe Gly Leu Gly His Asn Gly Met 355	360	365	1104	
gcc acc tac aat gcc gac gcc cgt ccg gac ttc tgg aag ctc caa gtc Ala Thr Tyr Asn Ala Asp Ala Arg Pro Asp Phe Trp Lys Leu Gln Val 370	375	380	1152	
acc acg act cgc aac gtc acg ggc gga cac ggt ttc ccc caa gcc ttt Thr Thr Thr Arg Asn Val Thr Gly Gly His Gly Phe Pro Gln Ala Phe 385	390	395	400	1200
gtc gac tgg ttc tgt ggt ggc ctc cag tac caa gtc gac cac cac tta Val Asp Trp Phe Cys Gly Leu Gln Tyr Gln Val Asp His His Leu 405	410	415	1248	
ttc ccc agc ctg ccc cga cac aat ctg gcc aag aca cac gca ctg gtc Phe Pro Ser Leu Pro Arg His Asn Leu Ala Lys Thr His Ala Leu Val 420	425	430	1296	
gaa tcg ttc tgc aag gag tgg ggt gtc cag tac cac gaa gcc gac ctt Glu Ser Phe Cys Lys Glu Trp Gly Val Gln Tyr His Glu Ala Asp Leu 435	440	445	1344	
gtg gac ggg acc atg gaa gtc ttg cac cat ttg ggc agc gtg gcc ggc Val Asp Gly Thr Met Glu Val Leu His His Leu Gly Ser Val Ala Gly 450	455	460	1392	
gaa ttc gtc gtg gat ttt gta cgc gat gga ccc gcc atg taa Glu Phe Val Val Asp Phe Val Arg Asp Gly Pro Ala Met 465	470	475	1434	
<210> 4				
<211> 477				
<212> PRT				
<213> Phaeodactylum tricor nutum				
<400> 4				
Met Gly Lys Gly Gly Asp Ala Arg Ala Ser Lys Gly Ser Thr Ala Ala 1 5 10 15				
Arg Lys Ile Ser Trp Gln Glu Val Lys Thr His Ala Ser Pro Glu Asp 20 25 30				
Ala Trp Ile Ile His Ser Asn Lys Val Tyr Asp Val Ser Asn Trp His 35 40 45				
Glu His Pro Gly Gly Ala Val Ile Phe Thr His Ala Gly Asp Asp Met 50 55 60				
Thr Asp Ile Phe Ala Ala Phe His Ala Pro Gly Ser Gln Ser Leu Met 65 70 75 80				

Lys Lys Phe Tyr Ile Gly Glu Leu Leu Pro Glu Thr Thr Gly Lys Glu
 85 90 95

 Pro Gln Gln Ile Ala Phe Glu Lys Gly Tyr Arg Asp Leu Arg Ser Lys
 100 105 110

 Leu Ile Met Met Gly Met Phe Lys Ser Asn Lys Trp Phe Tyr Val Tyr
 115 120 125

 Lys Cys Leu Ser Asn Met Ala Ile Trp Ala Ala Ala Cys Ala Leu Val
 130 135 140

 Phe Tyr Ser Asp Arg Phe Trp Val His Leu Ala Ser Ala Val Met Leu
 145 150 155 160

 Gly Thr Phe Phe Gln Gln Ser Gly Trp Leu Ala His Asp Phe Leu His
 165 170 175

 His Gln Val Phe Thr Lys Arg Lys His Gly Asp Leu Gly Leu Phe
 180 185 190

 Trp Gly Asn Leu Met Gln Gly Tyr Ser Val Gln Trp Trp Lys Asn Lys
 195 200 205

 His Asn Gly His His Ala Val Pro Asn Leu His Cys Ser Ser Ala Val
 210 215 220

 Ala Gln Asp Gly Asp Pro Asp Ile Asp Thr Met Pro Leu Leu Ala Trp
 225 230 235 240

 Ser Val Gln Gln Ala Gln Ser Tyr Arg Glu Leu Gln Ala Asp Gly Lys
 245 250 255

 Asp Ser Gly Leu Val Lys Phe Met Ile Arg Asn Gln Ser Tyr Phe Tyr
 260 265 270

 Phe Pro Ile Leu Leu Ala Arg Leu Ser Trp Leu Asn Glu Ser Phe
 275 280 285

 Lys Cys Ala Phe Gly Leu Gly Ala Ala Ser Glu Asn Ala Ala Leu Glu
 290 295 300

 Leu Lys Ala Lys Gly Leu Gln Tyr Pro Leu Leu Glu Lys Ala Gly Ile
 305 310 315 320

 Leu Leu His Tyr Ala Trp Met Leu Thr Val Ser Ser Gly Phe Gly Arg
 325 330 335

 Phe Ser Phe Ala Tyr Thr Ala Phe Tyr Phe Leu Thr Ala Thr Ala Ser
 340 345 350

 Cys Gly Phe Leu Leu Ala Ile Val Phe Gly Leu Gly His Asn Gly Met
 355 360 365

 Ala Thr Tyr Asn Ala Asp Ala Arg Pro Asp Phe Trp Lys Leu Gln Val
 370 375 380

Thr Thr Thr Arg Asn Val Thr Gly Gly His Gly Phe Pro Gln Ala Phe
 385 390 395 400

Val Asp Trp Phe Cys Gly Gly Leu Gln Tyr Gln Val Asp His His Leu
 405 410 415

Phe Pro Ser Leu Pro Arg His Asn Leu Ala Lys Thr His Ala Leu Val
 420 425 430

Glu Ser Phe Cys Lys Glu Trp Gly Val Gln Tyr His Glu Ala Asp Leu
 435 440 445

Val Asp Gly Thr Met Glu Val Leu His His Leu Gly Ser Val Ala Gly
 450 455 460

Glu Phe Val Val Asp Phe Val Arg Asp Gly Pro Ala Met
 465 470 475

<210> 5

<211> 1651

<212> DNA

<213> Phaeodactylum tricornutum

<220>

<221> CDS

<222> (67)..(1554)

<400> 5

gaagaaggaa catataaaaag taagccatct cctcggcacc atctaaagac ctaatatcta 60

ctcgtc atg gtt cgc ttt tca aca gcc gct cta ctt tct ctg tcg aca 108
 Met Val Arg Phe Ser Thr Ala Ala Leu Leu Ser Leu Ser Thr
 1 5 10

ttg aca act tca tgt att ggt gcc ttc cag ctg tct tcg cca gca caa 156
 Leu Thr Thr Ser Cys Ile Gly Ala Phe Gln Leu Ser Ser Pro Ala Gln
 15 20 25 30

ctt ccg aca agt agg ctt cgt cggt cat acg aac acg gcg ccg ctt tcg 204
 Leu Pro Thr Ser Arg Leu Arg Arg His Thr Asn Thr Ala Pro Leu Ser
 35 40 45

gcc gtg gcc gtc gac tcc ggt tct tcc gat ccg gcc ttg gta ggc aac 252
 Ala Val Ala Val Asp Ser Gly Ser Ser Asp Pro Ala Leu Val Gly Asn
 50 55 60

ctc ccc ctt ccc aac aac aat gat aat gag gac aag aac cgt aga atg 300
 Leu Pro Leu Pro Asn Asn Asn Asp Asn Glu Asp Lys Asn Arg Arg Met
 65 70 75

cca atg atg gac ttg aaa ggt att gct ctg tct ggt ctc aaa ggg caa 348
 Pro Met Met Asp Leu Lys Gly Ile Ala Leu Ser Gly Leu Lys Gly Gln
 80 85 90

gct ctt tcc gtc cga gcg gaa gat ttt cct cag gcg aaa gac ttg cgt 396
 Ala Leu Ser Val Arg Ala Glu Asp Phe Pro Gln Ala Lys Asp Leu Arg
 95 100 105 110

gcc gtc att ccg aaa gat tgc ttc gaa ccc gac acg gcc aaa tcg ttg		444
Ala Val Ile Pro Lys Asp Cys Phe Glu Pro Asp Thr Ala Lys Ser Leu		
115	120	125
gga tat ctt tcc gtt tca act atg ggg aca att ctc tgc tcc gtc gtc		492
Gly Tyr Leu Ser Val Ser Thr Met Gly Thr Ile Leu Cys Ser Val Val		
130	135	140
ggc gcg aac ctc ctt agt gtg ctc gat ccc tcc aat cca tta acc tgg		540
Gly Ala Asn Leu Leu Ser Val Leu Asp Pro Ser Asn Pro Leu Thr Trp		
145	150	155
cct ctc tgg gcc tac ggt gcc gtc acg ggg acg gtc gcc atg ggg		588
Pro Leu Trp Ala Ala Tyr Gly Ala Val Thr Gly Thr Val Ala Met Gly		
160	165	170
ctt tgg gtg ctg gcc cac gaa tgc gga cac ggc gcc ttt tcc aaa aac		636
Leu Trp Val Leu Ala His Glu Cys Gly His Gly Ala Phe Ser Lys Asn		
175	180	185
190		
cga tcc ctc cag gat gcc gtg ggg tac att atc cat tcc atc atg ctg		684
Arg Ser Leu Gln Asp Ala Val Gly Tyr Ile Ile His Ser Ile Met Leu		
195	200	205
gtg cca tac ttt agt tgg cag cga tcg cat gcc gtg cat cac cag tat		732
Val Pro Tyr Phe Ser Trp Gln Arg Ser His Ala Val His His Gln Tyr		
210	215	220
acc aat cat atg gaa ctg ggg gaa aca cac gtt cct gat cga gcc gat		780
Thr Asn His Met Glu Leu Gly Glu Thr His Val Pro Asp Arg Ala Asp		
225	230	235
aag gag ggc gag aag agc ctg gcg ctc cgc cag ttc atg ttg gat tcc		828
Lys Glu Gly Glu Lys Ser Leu Ala Leu Arg Gln Phe Met Leu Asp Ser		
240	245	250
ttt ggt aaa gac aag ggc atg aaa gca tac gga ggc ctc cag tcg ttt		876
Phe Gly Lys Asp Lys Gly Met Lys Ala Tyr Gly Leu Gln Ser Phe		
255	260	265
270		
ttg cat ctc atc gtg gga tgg cca gcc tac ctc ctg atc ggt gcg acc		924
Leu His Leu Ile Val Gly Trp Pro Ala Tyr Leu Leu Ile Gly Ala Thr		
275	280	285
ggt gga ccc gac cgt ggt atg acc aac cat ttt tat ccc aac cct ttg		972
Gly Gly Pro Asp Arg Gly Met Thr Asn His Phe Tyr Pro Asn Pro Leu		
290	295	300
tcg acg cca aca cag ccc aag aaa gaa ctt ttc cct ggg aac tgg aaa		1020
Ser Thr Pro Thr Gln Pro Lys Lys Glu Leu Phe Pro Gly Asn Trp Lys		
305	310	315
gaa aag gtc tac cag tca gat att gga atc gcc gcc gtt gtc ggc gcc		1068
Glu Lys Val Tyr Gln Ser Asp Ile Gly Ile Ala Ala Val Val Gly Ala		
320	325	330
ctc att gct tgg acc gcc act tcg ggt cta gcc ccc gtc atg gcc ttg		1116

11

Leu Ile Ala Trp Thr Ala Thr Ser Gly Leu Ala Pro Val Met Ala Leu 335	340	345	350	
tac ggt ggt ccc ttg atc gtc att aat gcc tgg ctg gta ctg tac acg Tyr Gly Gly Pro Leu Ile Val Ile Asn Ala Trp Leu Val Leu Tyr Thr 355	360	365		1164
tgg ttg caa cat aca gat acc gat gtt ccg cac ttt tcc tcc gac aac Trp Leu Gln His Thr Asp Thr Asp Val Pro His Phe Ser Ser Asp Asn 370	375	380		1212
cac aac ttt gtc aag ggc gca ctg cat acg atc gat cgt ccc tac gac His Asn Phe Val Lys Gly Ala Leu His Thr Ile Asp Arg Pro Tyr Asp 385	390	395		1260
aaa ctt gat ccc tgg gga atc ata gac ttt ctg cac cac aag att gga Lys Leu Asp Pro Trp Gly Ile Ile Asp Phe Leu His His Lys Ile Gly 400	405	410		1308
aca acg cat gtg gca cac cat ttt gac agt act atc ccc cac tat aag Thr Thr His Val Ala His His Phe Asp Ser Thr Ile Pro His Tyr Lys 415	420	425	430	1356
gct cag att gct acc gat gcc atc aaa gcc aag ttt cca gaa gtg tac Ala Gln Ile Ala Thr Asp Ala Ile Lys Ala Lys Phe Pro Glu Val Tyr 435	440	445		1404
ctc tat gac ccg aca cca att cca caa gcc atg tgg cgc gtc gcc aag Leu Tyr Asp Pro Thr Pro Ile Pro Gln Ala Met Trp Arg Val Ala Lys 450	455	460		1452
gga tgt act gca gta gag caa cgc ggt gac gcc tgg gtg tgg aaa aac Gly Cys Thr Ala Val Glu Gln Arg Gly Asp Ala Trp Val Trp Lys Asn 465	470	475		1500
gaa gga ata gaa gat ttg gtg gaa cat cgt caa agc aaa tta tcg agc Glu Gly Ile Glu Asp Leu Val Glu His Arg Gln Ser Lys Leu Ser Ser 480	485	490		1548
gaa taa agcaacatat cgctttatgg aagaacaaac gtccattgtg taaaaccctg Glu 495				1604
ataatttcaa tatttgttt tgttttaaaa aaaaaaaaaa aaaaaaaaaa 1651				
<210> 6				
<211> 495				
<212> PRT				
<213> Phaeodactylum tricor nutum				
<400> 5				
Met Val Arg Phe Ser Thr Ala Ala Leu Leu Ser Leu Ser Thr Leu Thr 1 5 10 15				
Thr Ser Cys Ile Gly Ala Phe Gln Leu Ser Ser Pro Ala Gln Leu Pro 20 25 30				

Thr Ser Arg Leu Arg Arg His Thr Asn Thr Ala Pro Leu Ser Ala Val
 35 40 45

Ala Val Asp Ser Gly Ser Ser Asp Pro Ala Leu Val Gly Asn Leu Pro
 50 55 60

Leu Pro Asn Asn Asn Asp Asn Glu Asp Lys Asn Arg Arg Met Pro Met
 65 70 75 80

Met Asp Leu Lys Gly Ile Ala Leu Ser Gly Leu Lys Gly Gln Ala Leu
 85 90 95

Ser Val Arg Ala Glu Asp Phe Pro Gln Ala Lys Asp Leu Arg Ala Val
 100 105 110

Ile Pro Lys Asp Cys Phe Glu Pro Asp Thr Ala Lys Ser Leu Gly Tyr
 115 120 125

Leu Ser Val Ser Thr Met Gly Thr Ile Leu Cys Ser Val Val Gly Ala
 130 135 140

Asn Leu Leu Ser Val Leu Asp Pro Ser Asn Pro Leu Thr Trp Pro Leu
 145 150 155 160

Trp Ala Ala Tyr Gly Ala Val Thr Gly Thr Val Ala Met Gly Leu Trp
 165 170 175

Val Leu Ala His Glu Cys Gly His Gly Ala Phe Ser Lys Asn Arg Ser
 180 185 190

Leu Gln Asp Ala Val Gly Tyr Ile Ile His Ser Ile Met Leu Val Pro
 195 200 205

Tyr Phe Ser Trp Gln Arg Ser His Ala Val His His Gln Tyr Thr Asn
 210 215 220

His Met Glu Leu Gly Glu Thr His Val Pro Asp Arg Ala Asp Lys Glu
 225 230 235 240

Gly Glu Lys Ser Leu Ala Leu Arg Gln Phe Met Leu Asp Ser Phe Gly
 245 250 255

Lys Asp Lys Gly Met Lys Ala Tyr Gly Gly Leu Gln Ser Phe Leu His
 260 265 270

Leu Ile Val Gly Trp Pro Ala Tyr Leu Leu Ile Gly Ala Thr Gly Gly
 275 280 285

Pro Asp Arg Gly Met Thr Asn His Phe Tyr Pro Asn Pro Leu Ser Thr
 290 295 300

Pro Thr Gln Pro Lys Lys Glu Leu Phe Pro Gly Asn Trp Lys Glu Lys
 305 310 315 320

Val Tyr Gln Ser Asp Ile Gly Ile Ala Ala Val Val Gly Ala Leu Ile
 325 330 335

Ala Trp Thr Ala Thr Ser Gly Leu Ala Pro Val Met Ala Leu Tyr Gly

340

345

350

Gly Pro Leu Ile Val Ile Asn Ala Trp Leu Val Leu Tyr Thr Trp Leu
355 360 365

Gln His Thr Asp Thr Asp Val Pro His Phe Ser Ser Asp Asn His Asn
370 375 380

Phe Val Lys Gly Ala Leu His Thr Ile Asp Arg Pro Tyr Asp Lys Leu
385 . 390 . 395 . 400

Asp Pro Trp Gly Ile Ile Asp Phe Leu His His Lys Ile Gly Thr Thr
405 410 415

His Val Ala His His Phe Asp Ser Thr Ile Pro His Tyr Lys Ala Gln
420 425 430

Ile Ala Thr Asp Ala Ile Lys Ala Lys Phe Pro Glu Val Tyr Leu Tyr
 435 440 445

Asp Pro Thr Pro Ile Pro Gln Ala Met Trp Arg Val Ala Lys Gly Cys
450 . . . 455 . . . 460

Thr Ala Val Glu Gln Arg Gly Asp Ala Trp Val Trp Lys Asn Glu Gly
465 470 475 480

Ile Glu Asp Leu Val Glu His Arg Gln Ser Lys Leu Ser Ser Glu
485 490 495

<210> 7
<211> 1578

<212> DNA

<213> Phys

<221>

<222> (1)

atg atg

Met Val Phe Ala Gly Gly Leu Gln Gln Gly Ser Leu Glu Glu Asn
1 5 10 15

atc gac gtc gag cac att gcc agt atg tct ctc ttc agc gac ttc ttc
 Ile Asp Val Glu His Ile Ala Ser Met Ser Leu Phe Ser Asp Phe Phe
 20 25 30

agt tat gtg tct tca act gtt ggt tcg tgg agc gta cac agt ata caa 144
 Ser Tyr Val Ser Ser Thr Val Gly Ser Trp Ser Val His Ser Ile Gln
 35 40 45

cct ttg aag cgc ctg acg agt aag aag cgt gtt tcg gaa agc gct gcc 192
Pro Leu Lys Arg Leu Thr Ser Lys Lys Arg Val Ser Glu Ser Ala Ala
50 55 60

gtg caa tgt ata tca gct gaa gtt cag aga aat tcg agt acc cag gga 240
Val Gln Cys Ile Ser Ala Glu Val Gln Arg Asn Ser Ser Thr Gln Gly

65

70

75

80

act gcg gag gca ctc gca gaa tca gtc gtg aag ccc acg aga cga agg Thr Ala Glu Ala Leu Ala Glu Ser Val Val Lys Pro Thr Arg Arg Arg	85	90	95	288	
tca tct cag tgg aag aag tcg aca cac ccc cta tca gaa gta gca gta Ser Ser Gln Trp Lys Lys Ser Thr His Pro Leu Ser Glu Val Ala Val	100	105	110	336	
cac aac aag cca agc gat tgc tgg att gtt gta aaa aac aag gtg tat His Asn Lys Pro Ser Asp Cys Trp Ile Val Val Lys Asn Lys Val Tyr	115	120	125	384	
gat gtt tcc aat ttt gcg gac gag cat ccc gga gga tca gtt att agt Asp Val Ser Asn Phe Ala Asp Glu His Pro Gly Gly Ser Val Ile Ser	130	135	140	432	
act tat ttt gga cga gac ggc aca gat gtt ttc tct agt ttt cat gca Thr Tyr Phe Gly Arg Asp Gly Thr Asp Val Phe Ser Ser Phe His Ala	145	150	155	160	480
gct tct aca tgg aaa att ctt caa gac ttt tac att ggt gac gtg gag Ala Ser Thr Trp Lys Ile Leu Gln Asp Phe Tyr Ile Gly Asp Val Glu	165	170	175	528	
agg gtg gag ccg act cca gag ctg ctg aaa gat ttc cga gaa atg aga Arg Val Glu Pro Thr Pro Glu Leu Leu Lys Asp Phe Arg Glu Met Arg	180	185	190	576	
gct ctt ttc ctg agg gag caa ctt ttc aaa agt tcg aaa ttg tac tat Ala Leu Phe Leu Arg Glu Gln Leu Phe Lys Ser Ser Lys Leu Tyr Tyr	195	200	205	624	
gtt atg aag ctg ctc acg aat gtt gct att ttt gct gcg agc att gca Val Met Lys Leu Leu Thr Asn Val Ala Ile Phe Ala Ala Ser Ile Ala	210	215	220	672	
ata ata tgt tgg agc aag act att tca gcg gtt ttg gct tca gct tgt Ile Ile Cys Trp Ser Lys Thr Ile Ser Ala Val Leu Ala Ser Ala Cys	225	230	235	240	720
atg atg gct ctg tgt ttc caa cag tgc gga tgg ctg tcc cat gat ttt Met Met Ala Leu Cys Phe Gln Gln Cys Gly Trp Leu Ser His Asp Phe	245	250	255	768	
ctc cac aat cag gtg ttt gag aca cgc tgg ctt aat gaa gtt gtc ggg Leu His Asn Gln Val Phe Glu Thr Arg Trp Leu Asn Glu Val Val Gly	260	265	270	816	
tat gtg atc ggc aac gcc gtt ctg ggg ttt agt aca ggg tgg tgg aag Tyr Val Ile Gly Asn Ala Val Leu Gly Phe Ser Thr Gly Trp Trp Lys	275	280	285	864	
gag aag cat aac ctt cat cat gct gct cca aat gaa tgc gat cag act Glu Lys His Asn Leu His His Ala Ala Pro Asn Glu Cys Asp Gln Thr	290	295	300	912	

tac caa cca att gat gaa gat att gat act ctc ccc ctc att gcc tgg Tyr Gln Pro Ile Asp Glu Asp Ile Asp Thr Leu Pro Leu Ile Ala Trp 305	310	315	320	960
agc aag gac ata ctg gcc aca gtt gag aat aag aca ttc ttg cga atc Ser Lys Asp Ile Leu Ala Thr Val Glu Asn Lys Thr Phe Leu Arg Ile 325	330	335		1008
ctc caa tac cag cat ctg ttc atg ggt ctg tta ttt ttc gcc cgt Leu Gln Tyr Gln His Leu Phe Phe Met Gly Leu Leu Phe Phe Ala Arg 340	345	350		1056
ggt agt tgg ctc ttt tgg agc tgg aga tat acc tct aca gca gtg ctc Gly Ser Trp Leu Phe Trp Ser Trp Arg Tyr Thr Ser Thr Ala Val Leu 355	360	365		1104
tca cct gtc gac agg ttg ttg gag aag gga act gtt ctg ttt cac tac Ser Pro Val Asp Arg Leu Leu Glu Lys Gly Thr Val Leu Phe His Tyr 370	375	380		1152
ttt tgg ttc gtc ggg aca gcg tgc tat ctt ctc cct ggt tgg aag cca Phe Trp Phe Val Gly Thr Ala Cys Tyr Leu Leu Pro Gly Trp Lys Pro 385	390	395	400	1200
tta gta tgg atg gcg gtg act gag ctc atg tcc ggc atg ctg ctg ggc Leu Val Trp Met Ala Val Thr Glu Leu Met Ser Gly Met Leu Leu Gly 405	410	415		1248
ttt gta ttt gta ctt agc cac aat ggg atg gag gtt tat aat tcg tct Phe Val Phe Val Leu Ser His Asn Gly Met Glu Val Tyr Asn Ser Ser 420	425	430		1296
aaa gaa ttc gtg agt gca cag atc gta tcc aca cgg gat atc aaa gga Lys Glu Phe Val Ser Ala Gln Ile Val Ser Thr Arg Asp Ile Lys Gly 435	440	445		1344
aac ata ttc aac gac tgg ttc act ggt ggc ctt aac agg caa ata gag Asn Ile Phe Asn Asp Trp Phe Thr Gly Gly Leu Asn Arg Gln Ile Glu 450	455	460		1392
cat cat ctt ttc cca aca atg ccc agg cat aat tta aac aaa ata gca His His Leu Phe Pro Thr Met Pro Arg His Asn Leu Asn Lys Ile Ala 465	470	475	480	1440
cct aga gtg gag gtg ttc tgt aag aaa cac ggt ctg gtg tac gaa gac Pro Arg Val Glu Val Phe Cys Lys Lys His Gly Leu Val Tyr Glu Asp 485	490	495		1488
gta tct att gct acc ggc act tgc aag gtt ttg aaa gca ttg aag gaa Val Ser Ile Ala Thr Gly Thr Cys Lys Val Leu Lys Ala Leu Lys Glu 500	505	510		1536
gtc gcg gag gct gcg gca gag cag cat gct acc acc agt taa Val Ala Glu Ala Ala Glu Gln His Ala Thr Thr Ser 515	520	525		1578

<211> 525

<212> PRT

<213> Physcomitrella patens

<400> 8

Met	Val	Phe	Ala	Gly	Gly	Gly	Leu	Gln	Gln	Gly	Ser	Leu	Glu	Glu	Asn
1				5				10					15		

Ile	Asp	Val	Glu	His	Ile	Ala	Ser	Met	Ser	Leu	Phe	Ser	Asp	Phe	Phe
					20				25				30		

Ser	Tyr	Val	Ser	Ser	Thr	Val	Gly	Ser	Trp	Ser	Val	His	Ser	Ile	Gln
					35			40				45			

Pro	Leu	Lys	Arg	Leu	Thr	Ser	Lys	Lys	Arg	Val	Ser	Glu	Ser	Ala	Ala
					50			55			60				

Val	Gln	Cys	Ile	Ser	Ala	Glu	Val	Gln	Arg	Asn	Ser	Ser	Thr	Gln	Gly
					65			70		75			80		

Thr	Ala	Glu	Ala	Leu	Ala	Glu	Ser	Val	Val	Lys	Pro	Thr	Arg	Arg	Arg
					85			90			95				

Ser	Ser	Gln	Trp	Lys	Lys	Ser	Thr	His	Pro	Leu	Ser	Glu	Val	Ala	Val
					100			105			110				

His	Asn	Lys	Pro	Ser	Asp	Cys	Trp	Ile	Val	Val	Lys	Asn	Lys	Val	Tyr
					115			120			125				

Asp	Val	Ser	Asn	Phe	Ala	Asp	Glu	His	Pro	Gly	Gly	Ser	Val	Ile	Ser
					130			135			140				

Thr	Tyr	Phe	Gly	Arg	Asp	Gly	Thr	Asp	Val	Phe	Ser	Ser	Phe	His	Ala
					145			150		155			160		

Ala	Ser	Thr	Trp	Lys	Ile	Leu	Gln	Asp	Phe	Tyr	Ile	Gly	Asp	Val	Glu
					165			170			175				

Arg	Val	Glu	Pro	Thr	Pro	Glu	Leu	Leu	Lys	Asp	Phe	Arg	Glu	Met	Arg
					180			185			190				

Ala	Leu	Phe	Leu	Arg	Glu	Gln	Leu	Phe	Lys	Ser	Ser	Lys	Leu	Tyr	Tyr
					195			200			205				

Val	Met	Lys	Leu	Leu	Thr	Asn	Val	Ala	Ile	Phe	Ala	Ala	Ser	Ile	Ala
					210			215			220				

Ile	Ile	Cys	Trp	Ser	Lys	Thr	Ile	Ser	Ala	Val	Leu	Ala	Ser	Ala	Cys
					225			230		235			240		

Met	Met	Ala	Leu	Cys	Phe	Gln	Gln	Cys	Gly	Trp	Leu	Ser	His	Asp	Phe
					245			250			255				

Leu	His	Asn	Gln	Val	Phe	Glu	Thr	Arg	Trp	Leu	Asn	Glu	Val	Val	Gly
					260			265			270				

Tyr	Val	Ile	Gly	Asn	Ala	Val	Leu	Gly	Phe	Ser	Thr	Gly	Trp	Trp	Lys
					275			280			285				

Glu Lys His Asn Leu His His Ala Ala Pro Asn Glu Cys Asp Gln Thr
 290 295 300
 Tyr Gln Pro Ile Asp Glu Asp Ile Asp Thr Leu Pro Leu Ile Ala Trp
 305 310 315 320
 Ser Lys Asp Ile Leu Ala Thr Val Glu Asn Lys Thr Phe Leu Arg Ile
 325 330 335
 Leu Gln Tyr Gln His Leu Phe Phe Met Gly Leu Leu Phe Phe Ala Arg
 340 345 350
 Gly Ser Trp Leu Phe Trp Ser Trp Arg Tyr Thr Ser Thr Ala Val Leu
 355 360 365
 Ser Pro Val Asp Arg Leu Leu Glu Lys Gly Thr Val Leu Phe His Tyr
 370 375 380
 Phe Trp Phe Val Gly Thr Ala Cys Tyr Leu Leu Pro Gly Trp Lys Pro
 385 390 395 400
 Leu Val Trp Met Ala Val Thr Glu Leu Met Ser Gly Met Leu Gly
 405 410 415
 Phe Val Phe Val Leu Ser His Asn Gly Met Glu Val Tyr Asn Ser Ser
 420 425 430
 Lys Glu Phe Val Ser Ala Gln Ile Val Ser Thr Arg Asp Ile Lys Gly
 435 440 445
 Asn Ile Phe Asn Asp Trp Phe Thr Gly Gly Leu Asn Arg Gln Ile Glu
 450 455 460
 His His Leu Phe Pro Thr Met Pro Arg His Asn Leu Asn Lys Ile Ala
 465 470 475 480
 Pro Arg Val Glu Val Phe Cys Lys Lys His Gly Leu Val Tyr Glu Asp
 485 490 495
 Val Ser Ile Ala Thr Gly Thr Cys Lys Val Leu Lys Ala Leu Lys Glu
 500 505 510
 Val Ala Glu Ala Ala Ala Glu Gln His Ala Thr Thr Ser
 515 520 525

<210> 9
 <211> 873
 <212> DNA
 <213> Physcomitrella patens

<220>
 <221> CDS
 <222> (1)...(873)

<400> 9
 atg gag gtc gtg gag aga ttc tac ggt gag ttg gat ggg aag gtc tcg 48

Met Glu Val Val Glu Arg Phe Tyr Gly Glu Leu Asp Gly Lys Val Ser			
1	5	10	15
cag ggc gtg aat gca ttg ctg ggt agt ttt ggg gtg gag ttg acg gat			96
Gln Gly Val Asn Ala Leu Leu Gly Ser Phe Gly Val Glu Leu Thr Asp			
20	25	30	
acg ccc act acc aaa ggc ttg ccc ctc gtt gac agt ccc aca ccc atc			144
Thr Pro Thr Thr Lys Gly Leu Pro Leu Val Asp Ser Pro Thr Pro Ile			
35	40	45	
gtc ctc ggt gtt tct gta tac ttg act att gtc att gga ggg ctt ttg			192
Val Leu Gly Val Ser Val Tyr Leu Thr Ile Val Ile Gly Gly Leu Leu			
50	55	60	
tgg ata aag gcc agg gat ctg aaa ccg cgc gcc tcg gag cca ttt ttg			240
Trp Ile Lys Ala Arg Asp Leu Lys Pro Arg Ala Ser Glu Pro Phe Leu			
65	70	75	80
ctc caa gct ttg gtg ctt gtg cac aac ctg ttc tgt ttt gcg ctc agt			288
Leu Gln Ala Leu Val Leu Val His Asn Leu Phe Cys Phe Ala Leu Ser			
85	90	95	
ctg tat atg tgc gtg ggc atc gct tat cag gct att acc tgg cgg tac			336
Leu Tyr Met Cys Val Gly Ile Ala Tyr Gln Ala Ile Thr Trp Arg Tyr			
100	105	110	
tct ctc tgg ggc aat gca tac aat cct aaa cat aaa gag atg gcg att			384
Ser Leu Trp Gly Asn Ala Tyr Asn Pro Lys His Lys Glu Met Ala Ile			
115	120	125	
ctg gta tac ttg ttc tac atg tct aag tac gtg gaa ttc atg gat acc			432
Leu Val Tyr Leu Phe Tyr Met Ser Lys Tyr Val Glu Phe Met Asp Thr			
130	135	140	
gtt atc atg ata ctg aag cgc agc acc agg caa ata agc ttc ctc cac			480
Val Ile Met Ile Leu Lys Arg Ser Thr Arg Gln Ile Ser Phe Leu His			
145	150	155	160
gtt tat cat cat tct tca att tcc ctc att tgg tgg gct att gct cat			528
Val Tyr His His Ser Ser Ile Ser Leu Ile Trp Trp Ala Ile Ala His			
165	170	175	
cac gct cct ggc ggt gaa gca tat tgg tct gcg gct ctg aac tca gga			576
His Ala Pro Gly Gly Glu Ala Tyr Trp Ser Ala Ala Leu Asn Ser Gly			
180	185	190	
gtg cat gtt ctc atg tat gcg tat tac ttc ttg gct gcc tgc ctt cga			624
Val His Val Leu Met Tyr Ala Tyr Tyr Phe Leu Ala Ala Cys Leu Arg			
195	200	205	
agt agc cca aag tta aaa aat aag tac ctt ttt tgg ggc agg tac ttg			672
Ser Ser Pro Lys Leu Lys Asn Lys Tyr Leu Phe Trp Gly Arg Tyr Leu			
210	215	220	
aca caa ttc caa atg ttc cag ttt atg ctg aac tta gtg cag gct tac			720
Thr Gln Phe Gln Met Phe Gln Phe Met Leu Asn Leu Val Gln Ala Tyr			
225	230	235	240

tac gac atg aaa acg aat gcg cca tat cca caa tgg ctg atc aag att		768	
Tyr Asp Met Lys Thr Asn Ala Pro Tyr Pro Gln Trp Leu Ile Lys Ile			
245	250	255	
ttg ttc tac tac atg atc tcg ttg ctg ttt ctt ttc ggc aat ttt tac		816	
Leu Phe Tyr Tyr Met Ile Ser Leu Leu Phe Leu Phe Gly Asn Phe Tyr			
260	265	270	
gta caa aaa tac atc aaa ccc tct gac gga aag caa aag gga gct aaa		864	
Val Gln Lys Tyr Ile Lys Pro Ser Asp Gly Lys Gln Gly Ala Lys			
275	280	285	
act gag tga		873	
Thr Glu			
290			
<210> 10			
<211> 290			
<212> PRT			
<213> Physcomitrella patens			
<400> 10			
Met Glu Val Val Glu Arg Phe Tyr Gly Glu Leu Asp Gly Lys Val Ser			
1	5	10	15
Gln Gly Val Asn Ala Leu Leu Gly Ser Phe Gly Val Glu Leu Thr Asp			
20	25	30	
Thr Pro Thr Thr Lys Gly Leu Pro Leu Val Asp Ser Pro Thr Pro Ile			
35	40	45	
Val Leu Gly Val Ser Val Tyr Leu Thr Ile Val Ile Gly Gly Leu Leu			
50	55	60	
Trp Ile Lys Ala Arg Asp Leu Lys Pro Arg Ala Ser Glu Pro Phe Leu			
65	70	75	80
Leu Gln Ala Leu Val Leu Val His Asn Leu Phe Cys Phe Ala Leu Ser			
85	90	95	
Leu Tyr Met Cys Val Gly Ile Ala Tyr Gln Ala Ile Thr Trp Arg Tyr			
100	105	110	
Ser Leu Trp Gly Asn Ala Tyr Asn Pro Lys His Lys Glu Met Ala Ile			
115	120	125	
Leu Val Tyr Leu Phe Tyr Met Ser Lys Tyr Val Glu Phe Met Asp Thr			
130	135	140	
Val Ile Met Ile Leu Lys Arg Ser Thr Arg Gln Ile Ser Phe Leu His			
145	150	155	160
Val Tyr His His Ser Ser Ile Ser Leu Ile Trp Trp Ala Ile Ala His			
165	170	175	
His Ala Pro Gly Gly Glu Ala Tyr Trp Ser Ala Ala Leu Asn Ser Gly			

20

180

185

190

Val His Val Leu Met Tyr Ala Tyr Tyr Phe Leu Ala Ala Cys Leu Arg
195 200 205

Ser Ser Pro Lys Leu Lys Asn Lys Tyr Leu Phe Trp Gly Arg Tyr Leu
210 215 220

Thr Gln Phe Gln Met Phe Gln Phe Met Leu Asn Leu Val Gln Ala Tyr
225 230 235 240

Tyr Asp Met Lys Thr Asn Ala Pro Tyr Pro Gln Trp Leu Ile Lys Ile
245 250 255

Leu Phe Tyr Tyr Met Ile Ser Leu Leu Phe Leu Phe Gly Asn Phe Tyr
260 265 270

Val Gln Lys Tyr Ile Lys Pro Ser Asp Gly Lys Gln Lys Gly Ala Lys
275 280 285

Thr Glu
290

<210> 11

<211> 1526

<212> DNA

<213> Phaeodactylum tricornutum

<220>

<221> CDS

<222> (92)..(1402)

<400> 11

gcttcgtta gcgtccata gtttgtaca cttggctgtg aaacgaatac gttcttggtc 60

tacttactac aacgaagcaa ccaccagcag c atg ggt aag gga ggt caa cga 112
Met Gly Lys Gly Gly Gln Arg
1 5

gct gta gct ccc aag agt gcc acc agc tct act ggc agt gct acc ctt 160
Ala Val Ala Pro Lys Ser Ala Thr Ser Ser Thr Gly Ser Ala Thr Leu
10 15 20

agc caa agc aag gaa cag gta tgg act tcg tcg tac aac cct ctg gcg 208
Ser Gln Ser Lys Glu Gln Val Trp Thr Ser Ser Tyr Asn Pro Leu Ala
25 30 35

aag gat tcc ccg gag ctg cca acc aaa ggc caa atc aag gcc gtc att 256
Lys Asp Ser Pro Glu Leu Pro Thr Lys Gly Gln Ile Lys Ala Val Ile
40 45 50 55

ccg aag gaa tgt ttc caa cgc tca gcc ttt tgg tct acc ttc tac ctg 304
Pro Lys Glu Cys Phe Gln Arg Ser Ala Phe Trp Ser Thr Phe Tyr Leu
60 65 70

atg cgc gat ctc gcc atg gct gcc ttt tgc tac gga acc tca cag 352
Met Arg Asp Leu Ala Met Ala Ala Phe Cys Tyr Gly Thr Ser Gln

75

80

85

gtc ctc tcc acc gac ctt ccc caa gac gcc acg ctc att ctg ccc tgg Val Leu Ser Thr Asp Leu Pro Gln Asp Ala Thr Leu Ile Leu Pro Trp	400
90 95 100	
gct ctc ggc tgg ggc gtc tac gcc ttt tgg atg gga acc att ctc acc Ala Leu Gly Trp Gly Val Tyr Ala Phe Trp Met Gly Thr Ile Leu Thr	448
105 110 115	
ggg cct tgg gta gtt gcg cac gaa tgt gga cac ggc gct tac tcc gac Gly Pro Trp Val Val Ala His Glu Cys Gly His Gly Ala Tyr Ser Asp	496
120 125 130 135	
tcc cag acg ttc aat gac gtg gtc ggc ttt atc gtc cac caa gct ttg Ser Gln Thr Phe Asn Asp Val Val Gly Phe Ile Val His Gln Ala Leu	544
140 145 150	
ctc gtc ccc tac ttt gcc tgg cag tac acc cac gcg aaa cac cac cgt Leu Val Pro Tyr Phe Ala Trp Gln Tyr Thr His Ala Lys His His Arg	592
155 160 165	
cga acc aac cat ctg gtg gac ggc gag tcc cac gtc cct tct acc gcc Arg Thr Asn His Leu Val Asp Gly Glu Ser His Val Pro Ser Thr Ala	640
170 175 180	
aag gat aac ggc ctc ggg ccg cac aac gag cga aac tcc ttc tac gcc Lys Asp Asn Gly Leu Gly Pro His Asn Glu Arg Asn Ser Phe Tyr Ala	688
185 190 195	
gcg tgg cac gag gcc atg gga gac ggc gcc ttt gcc gtc ttt caa gtc Ala Trp His Glu Ala Met Gly Asp Gly Ala Phe Ala Val Phe Gln Val	736
200 205 210 215	
tgg tcg cac ttg ttc gtc ggc tgg cct ctc tac ttg gcc ggt ctg gcc Trp Ser His Leu Phe Val Gly Trp Pro Leu Tyr Leu Ala Gly Leu Ala	784
220 225 230	
agt acc gga aag ctt gcg cac gaa ggt tgg tgg ctg gaa gaa cgg aac Ser Thr Gly Lys Leu Ala His Glu Gly Trp Trp Leu Glu Glu Arg Asn	832
235 240 245	
gcg att gcg gat cac ttt cga ccc agc tct ccc atg ttc ccc gcc aag Ala Ile Ala Asp His Phe Arg Pro Ser Ser Pro Met Phe Pro Ala Lys	880
250 255 260	
atc cgt gcc aag att gcc ctt tcc agc gcg acg gaa ctc gcc gtg ctc Ile Arg Ala Lys Ile Ala Leu Ser Ser Ala Thr Glu Leu Ala Val Leu	928
265 270 275	
gct gga ctc ttg tat gtc ggt aca cag gtc gga cac ctt ccc gtc ctg Ala Gly Leu Leu Tyr Val Gly Thr Gln Val Gly His Leu Pro Val Leu	976
280 285 290 295	
ctg tgg tac ttg gga ccg tac acc ttt gtc aac gct tgg ctt gta ctc Leu Trp Tyr Trp Gly Pro Tyr Thr Phe Val Asn Ala Trp Leu Val Leu	1024
300 305 310	

tac acg tgg ctg cag cat acg gac ccg tcc atc ccg cac tac ggt gaa Tyr Thr Trp Leu Gln His Thr Asp Pro Ser Ile Pro His Tyr Gly Glu 315 320 325	1072
ggc gag tgg acc tgg gtc aag ggc gcg ctc tct acc att gat cga gac Gly Glu Trp Thr Trp Val Lys Gly Ala Leu Ser Thr Ile Asp Arg Asp 330 335 340	1120
tac ggc atc ttc gat ttc ttt cac cac acc atc ggt tcc acg cac gtg Tyr Gly Ile Phe Asp Phe His His Thr Ile Gly Ser Thr His Val 345 350 355	1168
gta cac cat ttg ttc cac gaa atg ccc tgg tac aat gcc ggc att gcc Val His His Leu Phe His Glu Met Pro Trp Tyr Asn Ala Gly Ile Ala 360 365 370 375	1216
acg caa aag gtc aag gaa ttt ttg gaa ccc cag ggc ttg tac aat tac Thr Gln Lys Val Lys Glu Phe Leu Glu Pro Gln Gly Leu Tyr Asn Tyr 380 385 390	1264
gat ccg acc ccc tgg tac aag gcc atg tgg cgc att gcc cgg acc tgt Asp Pro Thr Pro Trp Tyr Lys Ala Met Trp Arg Ile Ala Arg Thr Cys 395 400 405	1312
cac tat gtg gag tca aac gag ggt gtg cag tat ttc aag agt atg gaa His Tyr Val Glu Ser Asn Glu Gly Val Gln Tyr Phe Lys Ser Met Glu 410 415 420	1360
aac gtg ccg ctg act aag gat gtg cga aac aaa gcc gca tga Asn Val Pro Leu Thr Lys Asp Val Arg Asn Lys Ala Ala 425 430 435	1402
aaaaaaagtgc caccgacgca taattttaca atcctaccaa caagaccaac attatatggt	1462
tttcgcttaa aagatagttt ttcttaccat ctgtgttagtc ggcacaaaaaaa aaaaaaaaaa	1522
aaaa	1526

<210> 12
<211> 436
<212> PRT
<213> Phaeodactylum tricornutum

Met Gly Lys Gly Gly Gln Arg Ala Val Ala Pro Lys Ser Ala Thr Ser 1 5 10 15
Ser Thr Gly Ser Ala Thr Leu Ser Gln Ser Lys Glu Gln Val Trp Thr 20 25 30
Ser Ser Tyr Asn Pro Leu Ala Lys Asp Ser Pro Glu Leu Pro Thr Lys 35 40 45
Gly Gln Ile Lys Ala Val Ile Pro Lys Glu Cys Phe Gln Arg Ser Ala 50 55 60
Phe Trp Ser Thr Phe Tyr Leu Met Arg Asp Leu Ala Met Ala Ala Ala

65	70	75	80
Phe Cys Tyr Gly Thr Ser Gln Val Leu Ser Thr Asp Leu Pro Gln Asp			
85	90	95	
Ala Thr Leu Ile Leu Pro Trp Ala Leu Gly Trp Gly Val Tyr Ala Phe			
100	105	110	
Trp Met Gly Thr Ile Leu Thr Gly Pro Trp Val Val Ala His Glu Cys			
115	120	125	
Gly His Gly Ala Tyr Ser Asp Ser Gln Thr Phe Asn Asp Val Val Gly			
130	135	140	
Phe Ile Val His Gln Ala Leu Leu Val Pro Tyr Phe Ala Trp Gln Tyr			
145	150	155	160
Thr His Ala Lys His His Arg Arg Thr Asn His Leu Val Asp Gly Glu			
165	170	175	
Ser His Val Pro Ser Thr Ala Lys Asp Asn Gly Leu Gly Pro His Asn			
180	185	190	
Glu Arg Asn Ser Phe Tyr Ala Ala Trp His Glu Ala Met Gly Asp Gly			
195	200	205	
Ala Phe Ala Val Phe Gln Val Trp Ser His Leu Phe Val Gly Trp Pro			
210	215	220	
Leu Tyr Leu Ala Gly Leu Ala Ser Thr Gly Lys Leu Ala His Glu Gly			
225	230	235	240
Trp Trp Leu Glu Glu Arg Asn Ala Ile Ala Asp His Phe Arg Pro Ser			
245	250	255	
Ser Pro Met Phe Pro Ala Lys Ile Arg Ala Lys Ile Ala Leu Ser Ser			
260	265	270	
Ala Thr Glu Leu Ala Val Leu Ala Gly Leu Leu Tyr Val Gly Thr Gln			
275	280	285	
Val Gly His Leu Pro Val Leu Leu Trp Tyr Trp Gly Pro Tyr Thr Phe			
290	295	300	
Val Asn Ala Trp Leu Val Leu Tyr Thr Trp Leu Gln His Thr Asp Pro			
305	310	315	320
Ser Ile Pro His Tyr Gly Glu Gly Trp Thr Trp Val Lys Gly Ala			
325	330	335	
Leu Ser Thr Ile Asp Arg Asp Tyr Gly Ile Phe Asp Phe Phe His His			
340	345	350	
Thr Ile Gly Ser Thr His Val Val His His Leu Phe His Glu Met Pro			
355	360	365	
Trp Tyr Asn Ala Gly Ile Ala Thr Gln Lys Val Lys Glu Phe Leu Glu			
370	375	380	

Pro Gln Gly Leu Tyr Asn Tyr Asp Pro Thr Pro Trp Tyr Lys Ala Met
 385 390 395 400

Trp Arg Ile Ala Arg Thr Cys His Tyr Val Glu Ser Asn Glu Gly Val
 405 410 415

Gln Tyr Phe Lys Ser Met Glu Asn Val Pro Leu Thr Lys Asp Val Arg
 420 425 430

Asn Lys Ala Ala
 435

<210> 13

<211> 3598

<212> DNA

<213> Unknown

<220>

<223> Sequenz stellt eine pflanzliche
 Promotor-Terminator-Expressionskassette in Vektor
 pUC19 dar

<400> 13

tcgcgcgttt cggtgatgac ggtaaaaacc tctgacacat gcagctcccg gagacggta 60

cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120

ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180

accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc 240

attcgccatt caggctgcgc aactgttggg aaggcgatc ggtgcgggcc tcttcgtat 300

tacgcccagct ggcgaaaggg ggatgtgctg caaggcgatt aagttggta acgcgcagggt 360

tttcccaagtc acgacgttgtt aaaacgacgg ccagtgaatt cggcgccgc agtcctcga 420

gcaaatttac acattgccac taaacgtcta aacccttgcata atttgtttt gttttactat 480

gtgtgttatg tatttgattt gcgataaatt ttatatttg gtactaaatt tataaacacct 540

tttatgctaa cgtttgccaa cacttagcaa ttgcgcaagtt gattaattga ttctaaatta 600

tttttgtctt ctaaatacat atactaatca actggaaatg taaatatttg ctaatatttc 660

tactatagga gaattaaagt gagtgaatat ggtaccacaa ggtttgaga tttaaattgtt 720

gcaatgctgc atggatggca tatacaccaa acattcaata attcttgagg ataataatgg 780

taccacacaa gatttgaggt gcatgaacgt cacgtggaca aaaggtttag taattttca 840

agacaacaat gttaccacac acaagtttg aggtgcacgc atggatgcc tggaaagt 900

ttaaaaatat ttggaaatg atttgcattgg aagccatgtg taaaaccatg acatccactt 960

ggaggatgca ataatgaaga aaactacaaa ttacatgca actagttatg catgtatct 1020

atataatgag gattttgcaa tactttcatt catacacact cactaagttt tacacgatta 1080
taatttccttc atagccagcc cacccgggtg ggccggccgc tgcaagtctag aaggcctcct 1140
gcttaatga gatatgcgag acgcctatga tcgcatacgata tttgcattca attctgttgt 1200
gcacgttgta aaaaacctga gcatgtgttag ctcagatcct taccggccgt ttccggttcat 1260
tctaataaatgat atatcacccg ttactatcgt atttttatga ataataattct ccgttcaatt 1320
tactgattgt ccgtcgacga attcgagctc ggcgcgccaa gcttggcgta atcatggtca 1380
tagctgttcc ctgtgtgaaa ttgttatccg ctcacaattc cacacaacat acgagccgga 1440
agcataaaagt gtaaaaggcctg gggtgctaa tgagttagt aactcacatt aattgcgttg 1500
cgctcaactgc ccgcatttcca gtcgggaaac ctgtcggtcc agctgcattta atgaatcgcc 1560
caacgcgcgg ggagagggcgg tttgcgttattt gggcgcttcc ccgcatttcgtc gtcactgac 1620
tcgctgcgtc cggtcggttcg gctgcggcga gcggtatcag ctcactaaaa ggccgttaata 1680
cggttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa agggcagcaa 1740
aaggccagga accgtaaaaaa ggccgcgttg ctggcgccccctt tccataggct ccgcggccct 1800
gacgagcattc acaaaaatcg acgctcaagt cagaggtggc gaaacccgac aggactataa 1860
agataaccagg cgtttcccccc tggaaagetcc ctcgtcgct ctcctgttcc gaccctgccc 1920
cttaccggat acctgtccgc ctttctccct tcgggaagcg tggcgcttcc tcatacgctca 1980
cgctgttaggt atctcagttc ggtgttaggtc gtcgcattca agctggctg tgtgcacgaa 2040
ccccccgttcc agcccgaccg ctgcgcctta tccggtaact atcgtcttga gtccaaacccg 2100
gtaagacacg acttatcgcc actggcagca gccactggta acaggattag cagagcgagg 2160
tatgttaggcg gtgcgtacaga gttcttgaag tgggtggctta actacggcta cactagaagg 2220
acagtatttg gtatctgcgc tctgtgttcc acgttttaccc tccggaaaaag agttggtagc 2280
tcttgatccg gcaaaacaaac caccgcgttcc agcgggtgtt tttttgttttgc caagcagcag 2340
attacgcgcga gaaaaaaaaagg atctcaagaa gatcctttga tcttttctac ggggtctgac 2400
gctcgttggaa acgaaaaactc acgttaaggg attttggtca tgagattatc aaaaaggatc 2460
ttcacctaga tcctttttaaa ttaaaaaatga agttttaaat caatctaaag tatatatgag 2520
taaacttgggt ctgacagtta ccaatgcctta atcagtggagg cacctatctc agcgatctgt 2580
ctatccatgtt catccatagt tgcctgactc cccgtcggtt agataactac gatacgggag 2640
ggcttaccat ctggccccag tgctgcaatg ataccgcgag accccacgctc accggctcc 2700
gatttatcag caataaaacca gccagccgga agggccgagc gcagaagtgg tcctgcaact 2760

ttatccgcct ccatccagtc tattaattgt tgccgggaag ctagagtaag tagtcgcca 2820
gttaatagtt tgcgcaacgt tggccatt gctacaggca tcgtggtgc acgctcgta 2880
tttggtatgg ctccattcag ctccggttcc caacgatcaa ggcgagttac atgatcccc 2940
atgttgtaaaaaaaagcggt tagtccttc ggtcctccga tcgttgtcag aagtaagttg 3000
gccgcagtgt tatcactcat ggttatggca gcactgcata attctcttac tgtcatgcca 3060
tccgtaagat gctttctgt gactggtgag tactcaacca agtcattctg agaáttagtgt 3120
atgcggcogac cgagttgctc ttgcccggcg tcaatacggg ataataccgc..gccacatagc 3180
agaactttää aagtgctcat cattggaaaaa cgtttctcgg ggcgaaaact ctcaaggatc 3240
ttaccgctgt tgagatccag ttcgatgtaa cccactcgta cacccaaactg atcttcagca 3300
tcttttactt tcaccagcgt ttctgggtga gcaaaaacag gaaggcaaaa tgccgcaaaa 3360
aagggaataaa gggcgacacg gaaatgttga atactcatac tcttccttt tcaatattat 3420
tgaagcattt atcagggtta ttgtctcatg agcggatáca tatttgaatg tatttagaaa 3480
aataaacaääa taggggttcc ggcgcacattt ccccgaaaag tgccacctga cgtctaagaa 3540
accattattt tcatgacatt aacctataaa aataggcgta tcacgaggcc ctttcgtc 3598

<210> 14

<211> 3590

<212> DNA

<213> Unknown

<220>

<223> Sequenz stellt eine pflanzliche Promotor-Terminator-Expressionskassette in Vektor pUC19 dar

<400> 14

tcgcgcgttt cggtgatgac ggtaaaaacc tctgacacat gcagctcccg gagacggtca 60

cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagccccgtg 120

ttggcgggtg tcggggctgg cttaaactatg cggcatcaga gcagattgta ctgagagtgc 180

accatatgcg gtgtcaaata ccccacagat gcgttaaggag aaaataccgc atcaggcgcc 240

attcggccatt caggctgcgc aactgttggg aagggcgate ggtgcgggcc tcttcgttat 300

tacggccagct ggcgaaaaggg ggatgtgtctg caaggcgatt aagtgggtt acggccagggt 360

tttccccatgc acgacgttgt aaaacgacgg ccagtgaatt cggcgccgg agctcctcga 420

gcaaatttac acattgccac taaaacgtcta aacccttgta atttgttttt gttttactat 480

gtgtgttatg tatttgattt gcgataaaatt tttatatttg gtactaaatt tataaacacct 540

tttatgctaa cgtttgccaa cacttagcaa tttgcaagtt gattaattga ttctaaatta 600
 ttttgtctt ctaaaatacat atactaatca actggaaatg taaatatttgc 660
 tactatagga gaattaaagt gagtgaatat ggtaccacaa ggtttggaga tttaattgtt 720
 gcaatgctgc atggatggca tatacaccaa acattcaata attcttgagg ataataatgg 780
 taccacacaa gatttgaggt gcatgaacgt cacgtggaca aaaggtagt taattttca 840
 agacaacaat gttaccacac acaagtttg aggtgcacgc atggatgccc tgtggaaagt 900
 ttaaaaatat ttggaaatg atttgcattt aagccatgtg taaaaccatg acatccactt 960
 ggaggatgca ataatgaaga aaactacaaa ttacatgca actagtatg catgtatct 1020
 atataatgag gatttgccaa tactttcatt catacadact cactaagtt tacacgatta 1080
 taatttcttc atagccagcg gatccgatata cgggccccct agcgtaacc ctgtttat 1140
 gagatatgctg agacgcctat gatcgcatga tatttgctt caattctgtt gtgcacgtt 1200
 taaaaaacct gagcatgtgt agctcagatc ctaccgccc gtttcggttc attctaata 1260
 atatatacacc cgttactatc gtattttat gaataatatt ctccgttcaa ttactgatt 1320
 gtccgtcgac gaattcgac tcggcgccaa aagcttggcg taatcatggt catagctgtt 1380
 tcctgtgtga aattgttatac cgctcacaat tccacacaac atacgagccg gaagcataaa 1440
 gtgtaaagcc tgggggtgcct aatgagttag ctaactcaca ttaattgcgt tgcgtcact 1500
 gcccgtttc cagtcgggaa acctgtcggt ccagctgcatt taatgaatcg gccaacgcgc 1560
 ggggagagggc ggtttcgat tggggcgctc ttccgcttcc tgcgtcactg actcgctgct 1620
 ctccgtcggtt cggctgcggc gagcggatc agctcactca aaggcggtaa tacggttatc 1680
 cacagaatca gggataacg caggaaagaa catgtgagca aaaggccagc aaaaggccag 1740
 gaaccgtaaa aaggccgcgt tgctggcgat tttccatagg ctccgccccct ctgacgagca 1800
 tcacaaaaat cgacgctcaa gtcagagggtg gcgaaaccccg acaggactat aaagatacca 1860
 ggcgtttccc cctggaaagct ccctcgatcg ctctcctgtt ccgaccctgc cgcttaccgg 1920
 atacctgtcc gcctttctcc ctccggaaag cgtggcgctt tctcatagct cacgctgttag 1980
 gtatctcagt tcgggtgttgc tgcgtcgatc caagctggc tgcgtgcacg aaccccccgt 2040
 tcagccccgac cgctgcgcct tatccggtaa ctatcgatc gaggccaaacc cggtaaagaca 2100
 cgacttatcg ccactggcag cagccactgg taacaggatt agcagagcga ggtatgttagg 2160
 cgggtctaca gagttcttgc agtgggtggcc taactacggc tacactagaa ggacagtatt 2220
 tggtatctgc gctctgatc agccagttac ctccggaaaa agagttggta gctcttgatc 2280

cggcaaaaca accaccgctg gtagcggtgg ttttttgtt tgcaagcagc agattacg 2340
 cagaaaaaaaaa ggatctcaag aagatcctt gatctttct acggggctcg acgctcagt 2400
 gaacgaaaac tcacgttaag ggattttgtt catgagatta tcaaaaagga tcttcaccta 2460
 gatcccttta aattaaaaat gaagtttaa atcaatctaa agtatataatg agtaaacttg 2520
 gtctgacagt taccaatgct taatcagtga ggcacctatc tcagcgatct gtctattcg 2580
 ttcatccata gttgcctgac tccccgtcgt gtagataact acgatacggg agggttacc 2640
 atctggcccc agtgctgcaa tgataccg 2700
 agcaataaac cagccagccg gaaggcccga gcgcagaagt ggtcctgcaa ctttatccgc 2760
 ctccatccag tctattaatt gttgccggga agctagagta agtagttcgc cagttatag 2820
 tttgcgcaac gttgttgcca ttgctacagg catcgtggc tcacgctcgt cgtttgtat 2880
 ggcttcattc agctccgggtt cccaaacgatc aaggcgagtt acatgatccc ccattgttg 2940
 caaaaaagcg gtttagctcct tcggcctcc gatcgttgca agaagtaagt tggccgcagt 3000
 gttatcactc atggttatgg cagcaactgca taattctttt actgtcatgc catccgtaa 3060
 atgctttct gtgactggcgt agtactcaac caagtcattc tgagaataatgt gtatgcggcg 3120
 accgagttgc tcttgcggcgtt cgtcaatacg ggataataacc gcgccacata gcagaacttt 3180
 aaaagtgcgc atcattggaa aacgttcttc gggcgaaaaa ctctcaagga tcttaccgct 3240
 gttgagatcc agttcgatgt aacccactcg tgcacccaaac tgatcttcag catctttac 3300
 tttcaccaggc gtttctgggt gagcaaaaac aggaaggcaa aatgcccgc 3360
 aaggggcgaca cggaaatgtt gaataactcat actcttcctt tttcaatatt attgaagcat 3420
 ttatcagggt tattgtctca tgagcggata catatggaa tgtatggaa aaaataaaca 3480
 aataggggtt ccgcgcacat ttccccgaaa agtgcacact gacgtctaag aaaccattat 3540
 tatcatgaca ttaacctata aaaataggcg tatcagcagg cccttcgtc 3590

<210> 15
 <211> 3584
 <212> DNA
 <213> Unknown

<220>
 <223> Sequenz stellt eine pflanzliche
 Promotor-Terminator-Expressionskassette in Vektor
 pUC19 dar

<400> 15
 tcgcgcgtt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60

cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
 ttggcgggtg tcggggctgg cttaactatg cgccatcaga gcagattgt a ctgagagtgc 180
 accatatgcg gtgtgaaata ccgcacagat gcgttaaggag aaaataccgc atcaggcgcc 240
 attcggcatt caggctgcgc aactgttggg aaggcgatc ggtgcggcc tcttcgtat 300
 tacgccagct ggcgaaaagg ggatgtgctg caaggcgatt aagttggta acgccagggt 360
 tttccagtc acgacgttgc aaaacgacgg ccagtgaatt cggcgcccg agtcctcga 420
 gcaaatttac acattgccac taaacgtcta aacccttgcattt gttttactat 480
 gtgtgttatg tatttgattt ggcataaatt ttatatttgc tataacacct 540
 ttatgctaa cgtttgc当地 cacttagcaa ttgcagtt gattaatttgc ttctaaatta 600
 ttttgc当地 ctaaatacat atactaatca actggaaatg taaatatttgc taaatatttgc 660
 tactatagga gaattaaagt gaggtaat ggtaccacaa ggtttggaga tttaatttgc 720
 gcaatgctgc atggatggca tatacaccaa acattcaata attttgagg ataataatgg 780
 taccacacaa gatttgagggt gcatgaacgt cacgtggaca aaaggtttag taattttca 840
 agacaacaat gttaccacac acaagtttgc aggtgc当地 atggatgccc tgtggaaagt 900
 taaaaatatttgc atttgc当地 aagccatgtg taaaaccatg acatccactt 960
 ggaggatgca ataataatgaa aaactacaaa ttacatgca actagttatg catgtatgc 1020
 atataatgag gattttgc当地 tactttgc当地 catacacact cactaagtt tacacgatta 1080
 taatttgc当地 atagccagca gatctggccgg catcgatccc gggccatggc ctgc当地 1140
 gagatatgc当地 agacgc当地 gatgc当地 tatttgc当地 caatttgc当地 gtgc当地 1200
 taaaaaacct gagcatgtg agctcagatc cttaccgc当地 gtttgc当地 atttgc当地 1260
 atataatcacc cgttactatc gtatatttgc当地 gaataatatttgc当地 tttactgatttgc当地 1320
 gtccgtcgac gagctcgccgg cgccaaagtttgc当地 ggc当地 tggc当地 tttccgtgt 1380
 gtgaaatttgc当地 tatccgctca caattccaca caacatacga gccc当地 aagca 1440
 agcctgggttgc当地 gcttgc当地 gcttgc当地 cacatttgc当地 gcttgc当地 cactgccc当地 1500
 tttccagtc当地 gggccatggc当地 cgttgc当地 gcttgc当地 gc当地 tttccgtgt 1560
 aggccgtttgc当地 cgttgc当地 gcttgc当地 tttccgtgt cactgactc当地 tggc当地 tttccgtgt 1620
 cttccggctg cggccagc当地 tttccgtgt cttccgtgt cttccgtgt cttccgtgt 1680
 atcaggggat aacgc当地 agaacatgtg agc当地 aaaaaggc cagc当地 aaaaaggc ccaggaaaccg 1740
 taaaaaggcc gcttgc当地 gcttgc当地 cttccgtgt cttccgtgt cttccgtgt 1800

aaatcgacgc tcaagtcaga ggtggcggaa cccgacagga ctataaagat accaggcggt 1860
tccccctgga agctccctcg tgccgctctcc tggccgacc ctggccctta ccggataacct 1920
gtcccgccctt ctcccatcgga gaagcgtggc gctttctcat agctcacgct gtaggtatct 1980
cagttcggtg taggtcggtc gctccaagct gggctgtgtg cacgaacccc cggttcagcc 2040
cgaccgctgc gccttatccg gtaactatcg tcttgagtcc aacccggtaa gacacgactt 2100
atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg taggcgggtc 2160
tacagagttc ttgaagtggt ggcctaacta cggtcacact agaaggacag tatttggtat 2220
ctgcgctctg ctgaagccag ttacctcgaa aaaaagagtt ggtagcttctt gatccggcaa 2280
acaaaccacc gctggtagcg gtggttttt tggccatggcag cagcagatta cgccgagaaa 2340
aaaaggatct caagaagatc ctgtatctt ttctacgggg tctgacgctc agtggAACGA 2400
aaactcacgt taaggattt tggcatgag attatcaaaaa aggtatctca cctagatcct 2460
tttaaattaa aaatgaagtt ttaaatcaat ctaaagtata tatgagtaaa cttggctga 2520
cagttaccaa tgcttaatca gtgaggcacc tatctcagcg atctgtctat ttctgtatc 2580
catagttgcc tgactccccg tcgtgttagat aactacgata cgggagggtt taccatctgg 2640
ccccagtgtc gcaatgatac cgccgagaccc acgtcacccg gctccagatt tatcagcaat 2700
aaaccagcca gccggaaggg ccgagcgcag aagtggctt gcaactttat ccgcctccat 2760
ccagtttattt aatttgttgcg gggaaagcttagt agtaagtagt tcgccagtttta atagtttgcg 2820
caacgttgcgatc gccattgtca caggcatcgat ggtgtcacgc tcgtcgatggc tgcgttgc 2880
attcagctcc ggttcccaac gatcaaggcg agttacatga tccccatgt tggcaaaaaa 2940
agcggttagc tccttcggtc ctccgatcgat tgcagaagt aagttggccg cagtgttac 3000
actcatggtt atggcagcac tgcataattc tcttactgtc atgcccattccg taagatgctt 3060
ttctgtgact ggtgagtagt caaccaagtc attctgagaa tagtgtatgc ggccgaccgag 3120
ttgctcttgc ccggcgtcaaa tacggataaa taccgcgcac catagcagaa ctttaaaagt 3180
gctcatcatt gaaaaacgtt ctccggcgaaaactctca aggtatcttccg cgctgttgc 3240
atccagttcg atgtaacccca ctcgtgcacc caactgatct tcagcatctt ttacttcac 3300
cagcgtttctt gggtagccaa aaacaggaag gcaaaatgcc gcaaaaaagg gaataaggc 3360
gacacggaaa tggtaatac tcatactttt ctttttcaaa tattattgaa gcatttatca 3420
gggttattgt ctcgtgagcg gatacatatt tgaatgtatt tagaaaaata aacaaatagg 3480
ggttccgcgc acatttcccc gaaaagtgc acctgacgatc taagaaacca ttatttatcat 3540

gacattaacc tataaaaata ggcgtatcac gaggcccttt cgtc 3584

<210> 16
<211> 4507
<212> DNA
<213> Unknown

<220>
<223> Sequenz stellt eine pflanzliche
Promotor-Terminator-Expressionskassette in Vektor
pUC19 dar

<400> 16
tcgcgcgttt cgggtatgac ggtgaaaacc tctgacacat gcagctcccg gagacggta 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc 240
attcgcatt caggctgcgc aactgttggg aagggcgatc ggtgcgggcc tcttcgctat 300
tacgccagct ggcgaaaggg ggatgtgctg caaggcgatt aagttggta acgccagggt 360
tttccagtc acgacgttgt aaaacgacgg ccagtgaatt cggcgcccg agtcctcga 420
gcaaatttac acattgccac taaacgtcta aacccttgta atttgtttt gtttactat 480
gtgtttatg tatttGattt ggcataaatt ttatatttg gtactaaatt tataacacct 540
tttatgctaa cgtttgccaa cacttagcaa ttgcaagtt gattaattga ttctaaatta 600
ttttgtctt ctaaatacat atactaatca actggaaatg taaatatttg ctaatatttc 660
tactatagga gaattaaagt gagtgaatat ggtaccacaa ggtttggaga tttaattgtt 720
gcaatgctgc atggatggca tatacaccaa acattcaata attctgagg ataataatgg 780
taccacacaa gatttggagggt gcatgaacgt cacgtggaca aaaggtttag taattttca 840
agacaacaat gttaccacac acaagtttg aggtgcacgc atggatgccc tgtggaaagt 900
ttaaaaatat ttggaaatg atttgcattgg aagccatgtg taaaaccatg acatccactt 960
ggaggatgca ataatgaaga aaactacaaa ttacatgca actagttatg catgtatct 1020
atataatgag gatattgcaa tacttcatt catacacact cactaagtt tacacgatta 1080
taatttcttc atagccagcc caccgcggtg ggcggccgccc tgcaatcttag aaggcctct 1140
gcttaatga gatatgcgag acgcctatga tcgcattata tttgcattca attctgttgt 1200
gcacgttgta aaaaacctga gcatgtgttag ctcagatcct taccgcgggt ttggatcat 1260
tctaataat atatcaccgg ttactatcgt attttatga ataatattct ccgttcaatt 1320

tactgattgt ccgtcgagca aatttacaca ttgccactaa acgtctaaac ccttgtaatt 1380
 tgttttgtt ttactatgtg tgttatgtat ttgatttgcg ataaatttt atatttggta 1440
 ctaaatttat aacacccccc atgctaacgt ttgccaacac ttagcaattt gcaagttgat 1500
 taatttgcattc taaaatttattt ttgtcttcta aatacatata ctaatcaact ggaaatgtaa 1560
 atatttgcata atatttctac tataggagaa ttaaagttag tgaatatggc accacaagg 1620
 ttggagattt aatttgttgcata atgctgcattt gatggcatat acaccaaaca ttcaataatt 1680
 cttgaggata ataatggtac cacacaagat ttgaggtgcata tgaacgtcac gtggacaaaa 1740
 ggtttagtaa ttttcaaga caacaatgtt accacacaca agttttgagg tgcatgcatt 1800
 gatgcctgt ggaaagtttta aaaatattt ggaaatgatt tgcatgaaacccatgtgtaa 1860
 aaccatgaca tccacttggaa ggttgcata atgaagaaaa ctacaaattt acatgcact 1920
 agttatgcattt gtagtctata taatgaggat ttgcataatc tttcatttc acacactcac 1980
 taagtttac acgattataa ttttttcata gccagcggat ccgatatcgg gcccgcgtc 2040
 gttaaccctg cttaatgag atatgcgaga cgccatgtat cgcatgatat ttgtttcaa 2100
 ttctgttgcac cacgttgcataaaaacccatgag catgtgtac tcagatcctt accgccccgtt 2160
 tcggttcatt ctaatgaata tatcaccgt tactatcgta tttttatgaa taatatttc 2220
 cgttcaattt actgattgtc cgtcgacgaa ttgcagctcg gcgcgcacccatgtaa 2280
 tcatggtcat agctgtttcc tttgtgaaat ttgtatccgc tcacaattcc acacaacata 2340
 cgagccggaa gcataaaagtgt taaaggctgg ggtgcctaattt ggttgcata actcacatta 2400
 attgcgttgc gtcactgccc cgctttccag tcggaaacc ttgcgtgcac gtcattaa 2460
 tgaatcgccc aacgcgcggg gagaggccgt ttgcgtattt ggcgccttc cgcttcctcg 2520
 ctcactgact cgctgcgcctc ggtcggttcgg ctgcggcgag cggatcagc tcactcaaag 2580
 gcggtaataac ggttatccac agaatcaggg gataacgcag gaaagaacat gtgagaaaa 2640
 ggccagcaaa aggccaggaa ccgtaaaaag gccgcgttgc tggcgccccccatggctc 2700
 cggcccccctg acgagcatca caaaaatcgatccatgcataagtc agaggtggcg aaacccgaca 2760
 ggactataaa gataccaggc gtttccccctt ggaagctccc tcgtgcgcctc tcctgttccg 2820
 accctgcgcgc ttaccggata cctgtccgc ttttcctt cgggaagcgt ggcgtttct 2880
 catacgatcacttcac gctgttaggtt tctcagttcg gttgttaggtcg ttgcgtccaa gctgggtgt 2940
 gtgcacgaac cccccgttca gcccgcacgc tgccgcctt cccgttaacta tcgtcttgc 3000
 tccaaacccgg taagacacga cttatcgccatccatgcagccactgttgc caggattagc 3060

agagcgaggt atgtaggcgg tgctacagag ttcttgaagt ggtggctaa ctacggctac 3120
 actagaagga cagtatgg tatctgcgct ctgctgaagc cagttacctt cgaaaaaaga 3180
 gttggtagct cttgatccgg caaacaaacc accgctggta gcggtggttt tttgtttgc 3240
 aagcagcaga ttacgcgcag aaaaaaagga tctcaagaag atcctttgat cttttctacg 3300
 gggtctgacg ctcagtgaa cggaaaactca cgtaaggga ttttggtcat gagattatca 3360
 aaaaggatct tcacctagat ccttttaaat taaaaatgaa gttttaaatc aatctaaagt 3420
 atatatgagt aaacttggtc tgacagttac caatgcttaa tcagtgaggc acctatctca 3480
 gogatctgtc tatttcgttc atccatagtt gcctgactcc ccgtcgtgta gataactacg 3540
 atacgggagg gcttaccatc tggccccagt gctgcaatga taccgcgaga cccacgctca 3600
 ccggctccag atttatcagc aataaaccag ccagccggaa gggccgagcg cagaagtgg 3660
 cctgcaactt tatccgcctc catccagtct attaattgtt gccgggaagc tagagtaagt 3720
 agttcgccag ttaatagttt gcgcaacgtt gttgccattg ctacaggcat cgtggtgtca 3780
 cgctcgtcgt ttggtatggc ttcattcagc tccggttccc aacgatcaaag gcgagttaca 3840
 tgateccccca tggtgcaaa aaaagcggtt agtccttcg gtccctcgat cggtgtcaga 3900
 agtaagttgg ccgcagtgtt atcactcatg gttatggcag cactgcataa ttctcttact 3960
 gtcatgccat ccgtaagatg ctttctgtg actgggtgagt actcaaccaa gtcattctga 4020
 gaatagtgtt tgccggcacc gagttgtct tgccggcgt caatacggga taataccgca 4080
 ccacatagca gaactttaaa agtgctcatc attggaaaac gttcttcggg gcgaaaactc 4140
 tcaaggatct taccgctgtt gagatccagt tcgatgtAAC ccactcgtgc acccaactga 4200
 tttcagcat ctttacttt caccagcggt tctgggtgag caaaaacagg aaggcaaaat 4260
 gcccggaaaa agggataaag ggacacacgg aaatgttggaa tactcataact cttccctttt 4320
 caatattatt gaagcatttA tcagggttat tgtctcatga gggatacat attgaatgt 4380
 atttagaaaa ataaacaaat aggggttccg cgacatttc cccgaaaagt gccacctgac 4440
 gtctaagaaa ccattattat catgacatttA acctataaaa ataggcgtat cacgaggccc 4500
 ttccgtc . 4507

<210> 17
 <211> 5410
 <212> DNA
 <213> Unknown
 <220>

<223> Sequenz stellt eine pflanzliche
Promotor-Terminator-Expressionskassette in Vektor
pUC19 dar

<400> 17

ttttgaaat gatttgcattt gaagccatgt gtaaaaaccat gacatccact tggaggatgc 60
 aataatgaag aaaactacaa atttacatgc aactagttat gcatgttagtc tatataatga 120
 ggattttgca atactttcat tcatacacac tcactaagtt ttacacgatt ataatttctt 180
 catagccagc ggatccgata tcggggccgc tagcgtaac cctgcttaa tgagatatgc 240
 gagacgccta tgatgcattt atatttgctt tcaattctgt tgtgcacgat gtaaaaaacc 300
 tgagcatgtg tagctcagat cttaccgc ggtttcggtt cattctaattt aatataatcac 360
 ccgttactat cgtatTTTta tgaataatat tctccgttca atttactgtat tgtccgtcga 420
 gcaaatttac acattgccac taaacgtcta aacccttgta atttggggattt gttttactat 480
 gtgtgttatg tatttgattt gcgataaaattt ttatatttg gtactaaattt tataacacct 540
 ttatgctaa cgtttgc当地 cacttagcaa ttgcagtt gattaatttga ttctaaattt 600
 ttttgc当地 ctaaatacat atactaatca actggaaatg taaatatttg ctaatatttc 660
 tactatagga gaattaaagt gagtgaatattt ggtaccacaa ggtttggaga tttaatttgc当地 720
 gcaatgctgc atggatggca tatacaccaaa acattcaata atttgc当地 gttttactat 780
 taccacacaa gatttgaggt gcatgaacgt cacgtggaca aaagggttttag taatTTTca 840
 agacaacaat gttaccacac acaagttttg aggtgcatttgc atggatgccc tgtggaaatg 900
 taaaaatatttgc当地 tttggaaatg atttgc当地 aagccatgtg taaaaccatg acatccactt 960
 ggaggatgca ataatgaaga aaactacaaa ttacatgc当地 actagttatg catgtgtct 1020
 atataatgag gattttgc当地 tacTTTcatt catacacact cactaagttt tacacgatta 1080
 taatttcttca atagccagca gatctgccgg cattcgatccc gggccatggc ctgcttaat 1140
 gagatatgc当地 agacgc当地 gatcgcatga tatttgctt caattctgtt gtgcacgat 1200
 taaaaaacctt gaggatgtgt agtcagatc ctaccgc当地 gtttc当地 atttcaatga 1260
 atataatcacc cgttactatc gtatTTTat gaataatattt ctccgttcaa ttactgatt 1320
 gtccgtcgac gagctcggcg cgccaaagctt ggcgtaatca tggcatagc tggatgtt 1380
 gtgaaattgt tatccgctca caattccaca caacatacga gccggaaagca taaagtgtaa 1440
 agcctgggt gcctaatgag tgagctaact cacattaattt gcttgc当地 cactgc当地 1500
 ttccagtc当地 gggaaacctgt cgtgccagct gcatatgc atcggccaac ggc当地 1560
 aggccggttgc cgtattggc gcttccgc ttccgtc当地 actgactc当地 tgc当地 1620

cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg gtaatacggt tatccacaga 1680
 atcagggat aacgcaggaa agaacatgtg agcaaaaaggc cagcaaaagg ccaggaaccg 1740
 taaaaaggcc gcgttgctgg cgttttcca taggctccgc cccccctgacg agcatcacaa 1800
 aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcggt 1860
 tccccctgga agctccctcg tgcgctctcc tggccgacc ctgcccgtta ccggataacct 1920
 gtccgccttt ctcccttcgg gaagcgtggc gctttctcat agctcacgct gtaggtatct 1980
 cagttcggtg taggtcgttc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc 2040
 cgaccgctgc gccttatccg gtaactatcg tctttagtcc aacccggtaa gacacgactt 2100
 atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg taggcggtgc 2160
 tacagagttc ttgaagtggt ggctctaacta cggtctacact agaaggacag tatttggtat 2220
 ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctttt gatccggcaa 2280
 acaaaccacc gctggtagcg gtggttttt tggccatcaag cagcagatta cgccgaaaa 2340
 aaaaggatct caagaagatc ctttgatott ttctacgggg tctgacgctc agtggAACGA 2400
 aaactcacgt taagggattt tggtcatgag attatcaaaa aggatcttca cctagatcct 2460
 tttaaattaa aaatgaagtt ttaaatcaat ctaaagtata tatgagtaaa ctggctga 2520
 cagttaccaa tgcttaatca gtgaggcacc tatctcagcg atctgtctat ttgcgttc 2580
 catagttgcc tgactcccg tcgtgttagat aactacgata cgggagggt taccatctgg 2640
 cccccagtgc gcaatgatac cgcgagaccc acgctcaccg gctccagatt tatcagcaat 2700
 aaaccagcca gccggaaggg ccgagcgcag aagtggctt gcaactttt ccgcctccat 2760
 ccagtctatt aatttgttgc gggaaatgtc agtaagtagt tcgcccgttta atagttgcg 2820
 caacgttgtt gccattgcta caggcatcgt ggtgtcacgc tcgtcggtt gtagggcttc 2880
 attcagctcc ggttcccaac gatcaaggcg agttacatga tccccatgt tgtcaaaaa 2940
 agcggtagc tccttcggtc ctccgatcgt tgcgttttgc gatggcttc 2980
 actcatggtt atggcagcac tgcataattc tcttactgtc atgcacatccg taagatgctt 3060
 ttctgtgact ggtgagttact caaccaagtc attctgagaa tagtgtatgc ggcgaccgag 3120
 ttgcgtttgc ccggcgtcaa tacggataa taccgcgcca catagcagaa cttaaaaagt 3180
 gctcatcatt ggaaaacgtt ctccggggcg .aaaactctca aggatcttac cgctgttgag 3240
 atccagttcg atgttaaccca ctcgtgcacc caactgtatct tcagcatctt ttacttcac 3300
 cagcgtttctt gggtagcaa aaacaggaag gcaaaaatgcc gcaaaaaagg gaataaggc 3360

gacacggaaa tggtgaatac tcatactttt ccttttcaa tattattgaa gcatttatca 3420
gggttattgt ctcatgagcg gatacatatt tgaatgtatt tagaaaaata aacaaatagg 3480
ggttccgcgc acatttcccc gaaaagtgcc acctgacgac taagaaacca ttattatcat 3540
gacattaacc tataaaaata ggcgtatcac gaggccctt cgtctcgcc gtttcggta 3600
tgacggtaaa aacctctgac acatgcagct cccggagacg gtcacagctt gtctgtaa 3660
ggatgccggg agcagacaag cccgtcaggg cgccgtcagcg ggtgttggcg ggtgtcgggg 3720
ctggcttaac tatgcggcat cagagoagat tgtactgaga gtgcaccata tgccgtgtga 3780
aataccgcac agatgcgtaa ggagaaaata ccgcattcagg cgccattcgc cattcaggct 3840
gaccaactgt tggaaagggc gatcggtgcg ggccctttcg ctattacgcc agctggcga 3900
agggggatgt gctgcaaggc gattaagttt ggttaacgcga gggttttccc agtcacgc 3960
ttgtaaaacg acggccagtg aattcggcgc gccgagctcc tcgagcaa attacacattt 4020
ccactaaacg totaaaccct tgtaattttt ttttggttt ctatgtgtt tatgtatgg 4080
atttgcata aatttttata tttggtaacta aatttataac acctttatg ctaacgttt 4140
ccaacactta gcaatttgca agttgattaa ttgattctaa attattttg tcttctaaat 4200
acatatacta atcaactgga aatgtaaata tttgctaata tttctactat aggagaatta 4260
aagttagtga atatggtacc acaaggttt gagatttat tggtaatg ctgcattggat 4320
ggcatataca ccaaacattc aataattttt gaggataata atggtaacc acaagatgg 4380
aggtgcattga acgtcacgtg gacaaaaggat ttagtaattt ttcaagacaa caatgttacc 4440
acacacaagt tttgaggtgc atgcattggat gccctgtgga aagtttaaaa atatttgg 4500
aatgatttgc atgaaaggcca tggtaaaac catgacatcc acttggagga tgcaataatg 4560
aagaaaaacta caaatttaca tgcaactagt tatgcattgtat gtcttatataa tgaggatgg 4620
gcaatactt cattcataca cactcactaa gtttacacg attataattt ctgcattggcc 4680
agcccaccgc ggtggggccgc cgccctgcagt ctggaaaggcc tccctgcttta atgagatatg 4740
cgagacgcct atgatgcatt gatatttgc ttcatttgc ttgtgcacgt tgtaaaaaac 4800
ctgagcatgt gtagctcaga tccttaccgc cggtttcgggt tcattctaat gaatatatca 4860
cccggtacta tcgtatggat atgaaataa ttctccgttc aatttactga ttgtccgtcg 4920
agcaaattta cacattgcca ctaaacgtct aaacccttgt aatttggatggatggatgg 4980
tgtgtttat gtatttgcatt tgcataat ttatattt ggtactaaat ttataacacc 5040
ttttatgcta acgtttgcca acacttagca atttgcattt tgattaaattt attctaaattt 5100

atttttgtct totaaataca tatactaattc aactggaaat gtaaatattt gctaatattt 5160
 ctactatagg agaattaaag tgagtgaata tggtaccaca aggtttggag atttaattgt 5220
 tgcaatgctg catggatggc atatacacca aacattcaat aattcttgag gataataatg 5280
 gtaccacaca agatitgagg tgcatgaacg tcacgtggac aaaaggttta gtaattttc 5340
 aagacaacaa tgttaccaca cacaagttt gaggtgcattt catggatgcc ctgtggaaag 5400
 tttaaaaata 5410

<210> 18
 <211> 648
 <212> DNA
 <213> Phaeodactylum tricornutum

<220>
 <221> CDS
 <222> (1)..(648)

<220>
 <223>

<400> 18
 tgg tgg aaa aac aag cac aac gga cac cac gcc gtc ccc aac ctc cac 48
 Trp Trp Lys Asn Lys His Asn Gly His His Ala Val Pro Asn Leu His
 1 5 10 15

tgc tcc tcc gca gtc gcg caa gat ggg gac ccg gac atc gat acc atg 96
 Cys Ser Ser Ala Val Ala Gln Asp Gly Asp Pro Asp Ile Asp Thr Met
 20 25 30

ccc ctt ctc gcc tgg tcc gtc cag caa gcc cag tct tac cgg gaa ctc 144
 Pro Leu Leu Ala Trp Ser Val Gln Gln Ala Gln Ser Tyr Arg Glu Leu
 35 40 45

caa gcc gac gga aag gat tcg ggt ttg gtc aag ttc atg atc cgt aac 192
 Gln Ala Asp Gly Lys Asp Ser Gly Leu Val Lys Phe Met Ile Arg Asn
 50 55 60

caa tcc tac ttt tac ttt ccc atc ttg ttg ctc gcc cgc ctg tcg tgg 240
 Gln Ser Tyr Phe Tyr Phe Pro Ile Leu Leu Leu Ala Arg Leu Ser Trp
 65 70 75 80

ttg aac gag tcc ttc aag tgc gcc ttt ggg ctt gga gct gcg tcg gag 288
 Leu Asn Glu Ser Phe Lys Cys Ala Phe Gly Leu Gly Ala Ala Ser Glu
 85 90 95

aac gct gct ctc gaa ctc aag gcc aag ggt ctt cag tac ccc ctt ttg 336
 Asn Ala Ala Leu Glu Leu Lys Ala Lys Gly Leu Gln Tyr Pro Leu Leu
 100 105 110

gaa aag gct ggc atc ctg ctg cac tac gct tgg atg ctt aca gtt tcg 384
 Glu Lys Ala Gly Ile Leu Leu His Tyr Ala Trp Met Leu Thr Val Ser
 115 120 125

tcc ggc ttt gga cgc ttc tcg ttc gcg tac acc gca ttt tac ttt cta 432
 Ser Gly Phe Gly Arg Phe Ser Phe Ala Tyr Thr Ala Phe Tyr Phe Leu
 130 135 140

acc gcg acc gcg tcc tgt gga ttc ttg ctc gcc att gtc ttt ggc ctc 480
 Thr Ala Thr Ala Ser Cys Gly Phe Leu Leu Ala Ile Val Phe Gly Leu
 145 150 155 160

ggc cac aac ggc atg gcc acc tac aat gcc gac gcc cgt ccg gac ttc 528
 Gly His Asn Gly Met Ala Thr Tyr Asn Ala Asp Ala Arg Pro Asp Phe
 165 170 175

tgg aag ctccaa gtc acc acg act cgc aac gtc acg ggc gga cac ggt 576
 Trp Lys Leu Gln Val Thr Thr Arg Asn Val Thr Gly Gly His Gly
 180 185 190

ttc ccc caa gcc ttt gtc gac tgg ttc tgt ggt ggc ctc cag tac caa 624
 Phe Pro Gln Ala Phe Val Asp Trp Phe Cys Gly Gly Leu Gln Tyr Gln
 195 200 205

gtc gac cac cac tta ttc ccc agc 648
 Val Asp His His Leu Phe Pro Ser
 210 215

<210> 19

<211> 216

<212> PRT

<213> Phaeodactylum tricornutum

<400> 19

Trp Trp Lys Asn Lys His Asn Gly His His Ala Val Pro Asn Leu His
 1 5 10 15

Cys Ser Ser Ala Val Ala Gln Asp Gly Asp Pro Asp Ile Asp Thr Met
 20 25 30

Pro Leu Leu Ala Trp Ser Val Gln Gln Ala Gln Ser Tyr Arg Glu Leu
 35 40 45

Gln Ala Asp Gly Lys Asp Ser Gly Leu Val Lys Phe Met Ile Arg Asn
 50 55 60

Gln Ser Tyr Phe Tyr Phe Pro Ile Leu Leu Ala Arg Leu Ser Trp
 65 70 75 80

Leu Asn Glu Ser Phe Lys Cys Ala Phe Gly Leu Gly Ala Ala Ser Glu
 85 90 95

Asn Ala Ala Leu Glu Leu Lys Ala Lys Gly Leu Gln Tyr Pro Leu Leu
 100 105 110

Glu Lys Ala Gly Ile Leu Leu His Tyr Ala Trp Met Leu Thr Val Ser
 115 120 125

Ser Gly Phe Gly Arg Phe Ser Phe Ala Tyr Thr Ala Phe Tyr Phe Leu
 130 135 140

Thr Ala Thr Ala Ser Cys Gly Phe Leu Leu Ala Ile Val Phe Gly Leu
 145 150 155 160

Gly His Asn Gly Met Ala Thr Tyr Asn Ala Asp Ala Arg Pro Asp Phe
 165 170 175

Trp Lys Leu Gln Val Thr Thr Arg Asn Val Thr Gly Gly His Gly
 180 185 190

Phe Pro Gln Ala Phe Val Asp Trp Phe Cys Gly Gly Leu Gln Tyr Gln
 195 200 205

Val Asp His His Leu Phe Pro Ser
 210 215

<210> 20

<211> 12093

<212> DNA

<213> Unknown

<220>

<223> pflanzlicher Expressionsvektor mit einer
 Promotor-Terminator-Expressionskassette

<400> 20

gatctggcgc cggccagcga gacgagcaag attggccgcc gccccaaacg atccgacagc 60

gcccggcagca caggtgcgca ggcaaattgc accaacgcac acagcgccag cagaatgcca 120

tagtggccgg tgacgtcggt cgagtgaacc agatcgca ggaggcccg cagcacccgc 180

ataatcaggc cgatgcccac agcgtcgagc gcgacagtgc tcagaattac gatcagggtt 240

atgttgggtt tcacgtctgg cctccggacc agcctccgtt ggtccgattt aacgcgcgga 300

ttcttttatca ctgataagtt ggtggacata ttatgtttat cagtgataaaa gtgtcaagca 360

tgacaaagtt gcagccgaat acagtgtacc gtgcggccct ggacctgttg aacgaggctg 420

gcgttagacgg tctgacgaca cgccaaactgg cggAACGGTT gggggttcag cagccggcgc 480

tttactggca cttcaggaac aagcggccgc tgctcgacgc actggccgaa gccatgctgg 540

cggagaatca tacgcattcg gtgcggagag ccgacgcacga ctggcgctca tttctgatcg 600

ggaatgcccgg cagtttcagg caggcgctgc tcgcctaccg cgatggccgc cgcatccatg 660

ccggcacgcg accggggcgca ccgcagatgg aaacggccga cgccgagtt cgcttcctct 720

gcgaggcggg tttttcggcc ggggacgccc tcaatgcgtt gatgacaatc agctacttca 780

ctgttggggc cgtgcttgag gagcaggccg cgcacagcga tgccggcagag cgccggcggca 840

ccgttgaaca ggctccgctc tcgcccgtgt tgccggccgc gatagacgcc ttgcacgaag 900

ccggtccgga cgcagcgttc gagcaggac tcgcgggtat tgtcgatgga ttggcgaaaa 960

ggaggcgtcg tgcgtcgac gttgaaggac cgagaaaggg tgacgattga tcaggaccgc 1020
tgccggagcg caacccactc actacagcag agccatgtag acaacatccc ctccccctt 1080
ccaccgogtc agacgcccgt agcagccgc tacgggcttt ttcatgcctt gccctagcgt 1140
ccaaggctca cggccgctc cggccctctt ggcggccttc tggcgcttcc cggccctctc 1200
gctcaactgac tcgctgcgtc cggcgttcg gctgcggcga gcggtatcag ctcactcaaa 1260
ggcggttaata cggttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa 1320
aggccagcaa aaggccagga accgtaaaaa ggccgcgttg ctggcgcccc tccataggct 1380
ccgccccctt gacgagcata acaaaaatcg acgctcaagt cagaggtggc gaaacccgac 1440
aggactataa agataccagg cgtttccccc tggaagctcc ctcgtgcgtc ctccctgtcc 1500
gaccctgccc cttaccggat acctgtccgc ctttctccct tggggaaagcg tggcgcttt 1560
ccgctgcata accctgcttc ggggtcatta tagcgatttt ttcggtatat ccattcttt 1620
tcgcacgata tacaggattt tgccaaaggg ttcgtgtaga ctttccttgg tgtatccaac 1680
ggcgtcagcc gggcaggata ggtgaagtag gcccacccgc gagcgggtgt tccttctca 1740
ctgtccctta ttgcacctg gcggtgctca acggaatcc tgctctgaga ggctggccgg 1800
ctaccgcccgg cgtaacagat gagggcaagc ggatggctga taaaaccaag ccaaccagga 1860
agggcagccc acctatcaag gtgtactgcc ttccagacga acgaagagcg attgaggaaa 1920
aggcggccggc ggcggcatg agcctgtcgg cctacctgct ggccgtcggc cagggctaca 1980
aaatcacggg cgtcggtggat tatgagcacg tccgcagact ggcccgcatc aatggcgacc 2040
tggggccctt gggcgccctg ctgaaactct ggctcaccga cgaccgcgc acggcgccgt 2100
tcggtgatgc cacatcctc gcccgtctgg cgaagatcga agagaagcag gacgagctt 2160
gcaagggtcat gatggcggtg gtccgcccga gggcagagcc atgactttt tagccgctaa 2220
aacggccggg gggcgccgt gattgccaag cacgtcccc tgcgctccat caagaagagc 2280
gacttcggcgg agctggtaaa gtacatcacc gacgagcaag gcaagaccga ggcgccttgc 2340
gacgctcacc gggctggttt ccctcgccgc tgggctggcg gccgtctatg gcccgtcaaa 2400
cgccgcagaa acggcgtcga agccgtgtgc gagacaccgc ggccgcggc gttgtggata 2460
cctcgccggaa aacttggccc tcactgacag atgagggcg gacgttgaca cttgaggggc 2520
cgactcaccc ggccggcgt tgacagatga gggcaggct cgatttcggc cggcgacgtg 2580
gagctggcca gcctcgcaaa tcggcgaaaa cgcctgattt tacgctgagtt tcccacagat 2640
gatgtggaca agcctgggaa taagtgcctt ggggtattga cacttgaggg gggcgactac 2700

tgacagatga ggggcgcgat ccttgacact tgagggcag agtgctgaca gatgagggc 2760
 gcacctattg acatttgagg ggctgtccac aggcaaaaa tccagcattt gcaagggttt 2820
 ccgcccgttt ttccggccacc gctaacctgt cttaaacctt gcttttaaac caatatttat 2880
 aaaccttgtt tttaaccagg gctgcgcctt gtgcgcgtga ccgcgcacgc cgaaggggg 2940
 tgccccccct tctcgaaaccc tcccgccccg ctaacgcggg cctcccatcc ccccaagggc 3000
 tgcgccttc ggcgcgaaac ggcctcaccc caaaaatggc agcgctggca gtccttgcac 3060
 ttgcgggat cggggcagta acgggatggg cgatcagccc gagcgcgcacg cccggaagca 3120
 ttgacgtgcc gcaggtgctg gcatcgacat tcagcgacca ggtgccggc agtgagggcg 3180
 cgggcctggg tggcggcctg cccttcaattt cggccgtcgg ggcattcacg gacttcatgg 3240
 cggggccggc aatttttacc ttggcatttcc ttggcatagt ggtcgccgggt gccgtgctcg 3300
 ttttcgggg tgcgataaaac ccagcgaacc attttaggtt ataggtaaga ttataccgag 3360
 gtatgaaaac gagaatttggc ctttacaga attactctat gaagcgcattt attaaaaag 3420
 ctaccaagac gaagaggatg aagaggatga ggaggcagat tgccttgcattt atattgacaa 3480
 tactgataag ataatatatac ttttatatag aagatatcgc cgtatgtt gatttcagg 3540
 ggcaaggcat aggcagcgcg cttatcaata tatctataga atggcacaag cataaaaaact 3600
 tgcattggact aatgcattgaa acccaggaca ataacctt atgcttgcattt ttctatcata 3660
 attggtaat gactccaact tattgatagt gttttatgtt cagataatgc ccgtatgactt 3720
 tgtcatgcag ctccaccgat ttggagaacg acagcgactt ccgtcccgac cgtgccaggt 3780
 gtcgcctcag attcaggat tggcgcctaa ttgcgtgcgtt atatcgatgtt ctgattacgt 3840
 gcagcttcc cttcaggcgg gattcataca gcccgcgcgc atccgtcattt catttcacca 3900
 cgtcaaaggc tgacagcagg ctcataagac gcccgcgtt cgcattatgtt cgttcaccga 3960
 atacgtgcgc aacaaccgtt ttccggagac tgcatacgc gtaaaaacgc cagcgctggc 4020
 gcgatttagc cccgacatag ccccaactgtt cgtccatttc cgcgcagacg atgacgtcac 4080
 tggccggctg tatgcgcgag gttaccgact gcccgcctgag ttttttaatgtt gacgtaaaat 4140
 cgtgttgagg ccaacgcggc taatgcgggc tggcgcctgg catccaaacgc catttcacca 4200
 cattatcaatg attttctgtt gcttaccggg ttggagaagcg gtgttaatgtt actgcagttt 4260
 ccatgttttta cggcagtgag agcagagata ggcgcgtatgtt ccggcgggtgc ttttgcgtt 4320
 acgcaccacc cctgcgttagt ctgaacagga gggacagctg atagacacag aagccactgg 4380
 agcaccctaa aaacaccatc atacactaaa tcagtaatgtt ggcagcatca cccataattt 4440

tggtttcaaa atcggctccg tcgataactat gttatacgcc aactttgaaa acaactttga 4500
 aaaagctgtt ttcttgttatt taagtttta gaatgcaagg aacagtgaat tggagttcgt 4560
 cttgttataa ttagcttctt gggtatctt taaaactgt agaaaagagg aaggaaataa 4620
 taaatggcta aaatgagaat atcacccgaa ttgaaaaaac tgatcgaaaa ataccgctgc 4680
 gtaaaagata cggaaggaat gtctcctgct aaggtatata agctggtggg agaaaatgaa 4740
 aacatatatt taaaaatgac ggacagccgg tataaaggga ccacctatga tgtggAACGG 4800
 gaaaaggaca tgatgctatg gctggaaagga aagctgcctg ttccaaaggt cctgcacttt 4860
 gaacggcatg atggctggag caatctgctc atgagtgagg ccgatggcgt cctttgctcg 4920
 gaagagtatg aagatgaaca aagccctgaa aagattatcg agctgtatgc ggagtgcac 4980
 aggcttttc actccatcgat catatcgat tgcctata cgaatagctt agacagccgc 5040
 ttagccgaat tggattactt actgaataac gatctggccg atgtggattt cgaaaactgg 5100
 gaagaagaca ctccatttaa agatccgcgc gagctgtatg attttttaaa gacggaaaag 5160
 cccgaagagg aacttgtctt ttcccacggc gacctgggag acagcaacat ctttgtaaa 5220
 gatggcaaag taagtggctt tattgatctt gggagaagcg gcagggccga caagtggat 5280
 gacattgcct tctgcgtccg gtcgatcagg gaggatatcg gggagaaca gatgtcgag 5340
 ctatTTTtgcctt acttactggg gatcaagcct gattgggaga aaataaaaata ttatattta 5400
 ctggatgaat tgTTTtagta cctagatgtg ggcacacat gcccggcaca agcaggagcg 5460
 caccgacttc ttccgcacatca agtgtttgg ctctcaggcc gaggcccacg gcaagtattt 5520
 gggcaagggg tcgtgttgat tcgtgcaggg caagattcg aataccaatg acgagaaggaa 5580
 cggccagacg gtctacggga ccgacttcat tgccgataag gtggattatc tggacaccaa 5640
 ggcaccaggc gggtaaaatc agaataagg gcacattgcc ccggcgtgag tcggggcaat 5700
 cccgcaagga gggtaatga atcggacgtt tgaccggaa gcatacaggc aagaactgat 5760
 cgacgcgggg tttccgccc aggatgccg aaccatcgca agccgcacccg tcatgcgtgc 5820
 gccccgcgaa accttccagt ccgtcgctc gatggtccag caagctacgg ccaagatcg 5880
 ggcgcacacg gtgcacactgg ctccccctgc cctgcccgcg ccacatggccg ccgtggagcg 5940
 ttgcgtcgt ctgcacacagg aggcggcagg tttggcgaag tcgtatgacca tcgcacacgc 6000
 aggaactatg acgaccaaga agcgaaaaac cggccggcag gacctggcaa aacaggtcag 6060
 cgaggccaaag caggccgcgt tgctgaaaca cacgaagcag cagatcaagg aaatgcagct 6120
 ttccctgttc gatattgcgc cgtggccggc cacgatgcga gcgatgcacaa acgacacggc 6180

ccgctctgcc ctgttcacca cgcgcaacaa gaaaatcccg cgcgaggcgc tgcaaaacaa 6240
ggtcatttc cacgtcaaca aggacgtgaa gatcacctac accggcgtcg agctgcggc 6300
cgacgatgac gaactggtgt ggcagcaggt gttggagtac gccaagcgc cccctatcgg 6360
cgagccgatc acttcacgt tctacgagct ttgcaggac ctgggctgt cgatcaatgg 6420
ccggtattac acgaaggccg aggaatgcct gtcgcgccta caggcgacgg cgatgggctt 6480
cacgtccgac cgcggtggc accttggaaatc ggtgtcgctg ctgcaccgt tccgcgtct 6540
ggaccgtggc aagaaaacgt cccgttgcca ggtcctgatc gacgaggaaa tcgtcgtgt 6600
gtttgctggc gaccactaca cgaattcat atgggagaag taccgcaagc tgcgcgcac 6660
ggccccgacgg atgttcgact atttcagctc gcaccggag ccgtacccgc tcaagctgg 6720
aaccttccgc ctcatgtgctg gatcgattc caccgcgtg aagaagtggc gcgagcaggt 6780
cgccgaagcc tgcaagagt tgcaaggcag cggcctggtg gaacacgcct gggtcaatga 6840
tgacctggtg cattgcaaac gctagggct tgcgggtca gttccggctg ggggttcagc 6900
agccagcgtt ttactggcat ttcaggaaca agcgggact gtcgacgca cttgcttcgc 6960
tcagtatcgc tcgggacgca cggcgcgctc tacgaactgc cgataaacag aggattaaaa 7020
ttgacaattt tgattaaggc tcaatttcga cggcttgag cggccgacgt gcaggatttc 7080
cgcgagatcc gattgtcgcc cctgaagaaaa gctccagaga tgcgggtca cgtttacgag 7140
cacgaggaga aaaagccat ggaggcggtc gctgaacggc tgcaagatgc cgtggcattc 7200
ggcgctaca tcgacggcga gatcattggg ctgtcggtct tcaaacagga ggacggcccc 7260
aaggacgctc acaaggcgca tctgtccggc gtttcgtgg agcccgaaca gcgaggccga 7320
ggggtcgccc gtatgctgt gccccgggtt cggccgggtt tattgctcgt gatgatcg 7380
cgacagattc caacggaaat ctgggtggatg cgcatcttca tccctggcgc acttaatatt 7440
tcgctattct ggagctgtt gtttatttcg gtctaccgc tgcggggcgg ggtcgccgc 7500
acggtaggcg ctgtgcagcc gctgatggc gtgttcatct ctggcgctct gctaggtac 7560
ccgatacgat tgcggcggt cctggggct atttgcggaa ctgcggcggt ggcgtgttg 7620
gtgttgcacac caaacgcagc gctagatcct gtcggcgctg cagcgggcct ggcggggcgc 7680
gtttccatgg ctgtcggtt cgtgctgacc cgcaagtggc aacctccgt gctctgc 7740
acctttacgg cctggcaact ggcggccggaa ggacttctgc tgcgtccagt agcttagtg 7800
tttgcgtccgc caatcccgat gcctacagga accaatgttc tccggcgtggc gtcggcg 7860
ctgatcgccgag cgggtttaac ctacttcctt tgggttccggg ggatctcgac actcgaaacct 7920

acagttgttt ctttactggg ctttctcagc cccagatctg gggtcgatca gccggggatg 7980
catcaggccg acagtcggaa ctgcgggtcc ccgacctgta ccattcggtg agcaatggat 8040
aggggagttg atatcgtaa cgttcaattc taaagaaata gcgccactca gcttcctcag 8100
cggtttatc cagcgatttc ctattatgtc ggcatagttc tcaagatcga cagcctgtca 8160
cggttaagcg agaaaatgaat aagaaggctg ataattcggta tctctgcgag ggagatgata 8220
tttgatcaca ggcagcaacg ctctgtcattc gttacaatca acatgctacc ctccgcgaga 8280
tcatccgtgt ttcaaaccgg gcagcttagt tgccgttctt ccgaatacgca tcggtaacat 8340
gagcaaagtc tgccgccta caacggctct cccgctgacg ccgtcccgga ctgatgggct 8400
gcctgtatcg agtgggtgatt ttgtgcccgg ctgcccgtcg gggagctgtt ggctggctgg 8460
tggcaggata tattgtggtg taaacaaatt gacgcttaga caacttaata acacattgcg 8520
gacgtttta atgtactggg gtgggttttc tttcaccag tgagacgggc aacagctgat 8580
tgcccttcac cgccctggccc tgagagagtt gcagcaagcg gtccacgctg gttgccccca 8640
gcaggcgaaa atccctgttttgc atgggtggttc cgaaatcgcc aaaatccctt ataaatcaaa 8700
agaatagccc gagatagggt tgagtgttgc tccagtttgg aacaagatcg cactattaaa 8760
gaacgtggac tccaacgtca aagggcgaaa aaccgtctat cagggcgatg gcccactacg 8820
tgaaccatca cccaaatcaa gttttttggg gtcgaggtgc cgtaaagcac taaatcgaa 8880
ccctaaaggg agccccccgat ttagagcttgc acggggaaag ccggcgaacg tggcgagaaa 8940
ggaagggaag aaagcgaaag gagcgggcgc cattcaggct ggcgaactgt tgggaaggc 9000
gatcggtgcg .ggccctttcg ctattacgccc agctggcgaa agggggatgt gctgcaaggc 9060
gattaagttg ggtAACGCCA gggttttccc agtcacgacg ttgtaaaacg acggccagtgc 9120
aattaattcc catcttggaaa gaaatatagt ttaaatatattt attgataaaaa taacaagtca 9180
ggtattatag tccaagcaaa aacataaatt tattgatgca agtttaaatt cagaaatatt 9240
tcaataactg attatatcag ctggtagcatt gccgttagatg, aaagactgag tgcgatatta 9300
tgtgtataac ataaattgtat gatatacgta gcttagctca tcgggggatc cgtcgaagct 9360
agcttgggtc ccgctcagaa gaactcgta agaaggcgat agaaggcgat ggcgtgcgaa 9420
tcgggagcgg cgataccgta aagcacgagg aagcggtcag cccattcgcc gccaagctct 9480
tcagcaatat cacgggttagc caacgctatg tcctgtatgc ggtccgccc acccagccgg 9540
ccacagtcga tgaatccaga aaagcgccca tttccacca tgatattcgg caagcaggca 9600
tcgcccattggg tcacgacgag atcctcgccg tcgggcatgc ggcgcattgag cctggcgaac 9660

agttcggtcg ggcggccccc ctgatgtct tcgtccagat catcctgatc gacaagaccg 9720
 gcttccatcc gagtaacgtgc tcgctcgatg cgatgttcg cttgggtggc gaatggcag 9780
 gtagccggat caagcgtatg cagccgccgc attgcattcag ccatgatgga tactttctcg 9840
 gcaggagcaa ggtgagatga caggagatcc tgccccggca cttcgcccaa tagcagccag 9900
 tcccttcccg ctcaagtgc aacgtcgagc acagctgcgc aaggaacgcc cgctgtggcc 9960
 agccacgata gcccgctgc ctcgtctgc agttcattca gggcacggga caggtcggtc 10020
 ttgacaaaaaa gaaccggcgccc cccctgcgct gacagccggaa acacggccgc atcagagcag 10080
 ccgattgtct gtgtgcccc gtcatacgaa aatagcccttc ccacccaagc ggccggagaa 10140
 cctgcgtgca atccatcttg ttcaatccaa gctccatgg gccctcgact agagtcgaga 10200
 tctggattga gagtgaatat gagactctaa ttggataccg agggaaattt atgaaacgtc 10260
 agtggagcat tttgacaag aaatatttgc tagctgatag tgacctttagg cgactttga 10320
 acgcgcaata atggtttctg acgtatgtgc ttagctcatt aaactccaga aaccggccgc 10380
 tgagtggctc cttcaacgtt gcggttctgt cagttccaaa cgtaaaacgg cttgtccgc 10440
 gtcatcgccg ggggtcataaa cgtgactccc ttaattctcc gctcatgatc ttgatcccct 10500
 gcccacatcag atccttggcg gcaagaaagc catccagttt actttgcagg gcttcccaac 10560
 cttaccagag ggcggcccgag ctggcaattc cggttcgtt gctgtccatia aaaccggccca 10620
 gtctagctat cgccatgtaa gcccactgca agtacactgc tttctcttg cgcttgcgtt 10680
 ttcccttgc cagatagccc agtagctgac attcatccgg ggtcagcacc gtttctgcgg 10740
 actggcttc tacgtgttcc gcttcctta gcaagcccttg cgccctgagt gcttgcggca 10800
 gctgtgaagct tgcatgcctg caggtcgacg ggcggcccgag ctccctcgagc aaatttacac 10860
 attgccacta aacgtctaaa cccttgtat ttgttttgtt tttactatgt gtgttatgt 10920
 ttgttgc gataaatttt tatatttggt actaaattta taacacccctt tatgctaacg 10980
 ttgtccaca ctttagcaatt tgcaagttga ttaattgatt ctaaatttattt tttgtcttct 11040
 aaatacatat actaatcaac tggaaatgtaa aatatttgc aatatttcta ctataggaga 11100
 attaaagtga gtgaatatgg taccacaagg tttggagatt taattgtgc aatgctgcat 11160
 ggatggcata tacacccaaac attcaataat tcttgaggat aataatggta ccacacaaga 11220
 tttgaggtgc atgaacgtca cgtggacaaa aggttttagta attttcaag acaacaatgt 11280
 taccacacac aagtttgag gtgcattgc ggtatgcctg tggaaagttt aaaaatattt 11340
 tggaaatgtat ttgcattggaa gccatgtgtaa aaccatgac atccacttgg aggatgcaat 11400

aatgaagaaa actacaaatt tacatgcaac tagttatgca ttagtctat ataatgagga 11460
 ttttgcata ctttcattca tacacactca ctaagttta cacgattata atttcttcat 11520
 agccagccca ccgcgggtggg cggccgcctg cagtctagaa ggcctcctgc tttaatgaga 11580
 tatgcgagac gcctatgatc gcatgatatt tgcttcaat tctgttgtgc acgttgaaa 11640
 aaacctgagc atgtgttagct cagatccta ccgcgggttt cggttcatc taatgaatat 11700
 atcacccgtt actatcgat ttttatgaat aatattctcc gttcaattta ctgattgtcc 11760
 gtcgacgaat tcgagctcg cgccgcctca gaggatcgat gaattcagat cggctgagtg 11820
 gctcctcaa cgttgcgtt ctgtcagtt caaacgtaaa acggcttgc ccgcgtcattc 11880
 ggccccggtc ataacgtgac tcccttaatt ctccgctcat gatcagattg tcgttccc 11940
 cttcagttt aaactatcag tgtttgacag gatattttgg cgggtaaacc taagagaaaa 12000
 gagcgtttat tagaataatc ggatatttaa aaggcgtga aaaggtttat cttcgtcca 12060
 tttgtatgtg catgccaacc acagggttcc cca 12093

<210> 21

<211> 12085

<212> DNA

<213> Unknown

<220>

<223> pflanzlicher Expressionsvektor mit einer Promotor-Terminator-Expressionskassette.

<400> 21

gatctggcgc cggccagcga gacgagcaag attggccgcc gcccggaaacg atccgacagc 60
 gcgccccagca caggtgcgcga ggcaaaattgc accaacgcattt acagcgcagg cagaatgc 120
 tagtggccgg tgacgtcggtt cgagtgaacc agatgcgcga ggaggcccg cagcacccggc 180
 ataatcaggc cgatgcccac agcgtcgaggc ggcacagtgc tcagaattttt gatcagggtt 240
 atgttgggtt tcacgtctgg cctccggacc agcctccgct ggtccgattt aacgcgcgga 300
 ttctttatca ctgatáagtt ggtggacata ttatgtttat cagtataaaa gtgtcaagca 360
 tgacaaagtt gcagccgaat acagtgtatcc gtgccgcct ggacctgttgc aacgaggctg 420
 gcgttagacgg tctgacgaca cgcaactgg cggAACGGTTT ggggggttcag cagccggcgc 480
 ttactggca cttcaggaac aagcggccgc tgctcgacgc actggccgaa gccatgctgg 540
 cggagaatca tacgcattcg gtgccgagag ccgacgacga ctggcgctca ttctgtatcg 600
 ggaatgcccgg cagcttcagg caggcgctgc tgccttaccg cgatggccgcg cgcattccatg 660

ccggcacgcg accgggcgca ccgcagatgg aaacggccga cgcgagctt cgcttcctct 720
 gcgaggcggg ttttcggcc ggggacgccc tcaatgcgct gatgacaatc agctacttca 780
 ctgttggggc cgtgcttgag gagcaggccg gcgcacagcga tgccggcgag cgccggcgca 840
 ccgttgaaca ggctccgctc tcgcccgtgt tgccggccgc gatagacgcc ttgcacgaag 900
 ccggtccgga cgcagcgttc gagcaggac tcgcggtgat tgtcgatgga ttggcgaaaa 960
 ggaggctcggt tgtcaggaac gttgaaggac cgagaaaggg tgacgattga tcaggaccgc 1020
 tgccggagcg caacccactc actacagcag agccatgtag acaacatccc ctcccccttt 1080
 ccaccgcgtc agacgcccgt agcagccgc tacggcttt ttcatgcctt gccctagcgt 1140
 ccaaggctca cggccgcgtc cggcctctct ggcggccttc tggcgctctt ccgccttcctc 1200
 gctcaactgac tcgctgcgtc cggtcgttcg gctgcggcga gcggtatcag ctcaactaaa 1260
 ggcggtaata cggtttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa 1320
 aggccagcaa aaggccagga accgtaaaaaa ggccgcgttg ctggcgtttt tccataggct 1380
 ccgccttcgtc gacgagcatc acaaaaatcg acgctcaagt cagaggtggc gaaacccgac 1440
 aggactataa agataaccagg cgtttcccccc tggaaagctcc ctgcgtgcgt ctccctgttcc 1500
 gaccctgccc cttaccggat acctgtccgc ctttctccct tcgggaagcg tggcgctttt 1560
 ccgctgcata accctgcttc ggggtcatta tagcgtatcc ttccgtatata ccattcccttt 1620
 tcgcacgata tacaggattt tgccaaaggg ttgcgtgtaga ctttccttgg tgtatccaaac 1680
 ggcgtcagcc gggcaggata ggtgaagtag gcccacccgc gagcgggtgt tccttcttca 1740
 ctgtccctta ttgcgcacctg gcggtgctca acggaaatcc tgctctgcga ggctggccgg 1800
 ctaccgcggc cgtaacagat gagggcaagc ggtggctga tgaaaccaag ccaaccagga 1860
 agggcagccc acctatcaag gtgtactgcc ttccagacga acgaagagcg attgaggaaa 1920
 aggcggccggc ggccggcatg agcctgtcgg cctacctgct ggccgtcggc cagggttaca 1980
 aaatcacggg cgtcgtggac tatgagcacc tccgcgagct ggcccgcatc aatggcgacc 2040
 tggccgcctt gggcggtctg ctgaaactct ggctcaccga cgacccgcgc acggcgcgg 2100
 tcggtgatgc cacgatcctc gccctgctgg cgaagatcga agagaagcag gacgagcttgc 2160
 gcaaggatcat gatggcggtg gtccggccga gggcagagcc atgactttt tagccgctaa 2220
 aacggccggg gggcgtcggt gattgccaag cacgtccccca tgcgctccat caagaagagc 2280
 gacttcgcgg agctggtgaa gtacatcacc gacgagcaag gcaagaccga ggcgccttgc 2340
 gacgctcacc gggctgggttg ccctcgccgc tgggcgtggcg gccgtctatg gccctgcaaa 2400

cgcgccagaa acgcccgtcga agccgtgtgc gagacaccgc ggccgcggc gttgtggata 2460
 cctcgccaa aacttggccc tcactgacag atgaggggag gacgttgaca cttgaggggc 2520
 cgactcaccc ggcgcggcgt tgacagatga ggggcaggct cgatttcggc cggcgacgtg 2580
 gagctggcca gcctcgcaaa tcggcgaaaa cgcctgattt tacgcgagtt tcccacagat 2640
 gatgtggaca agcctggggta taagtgcacct gcggttattga cacttgaggg gcgcgactac 2700
 tgacagatga ggggcgcgat ctttgacact tgagggcag agtgctgaca gatgaggggc 2760
 gcaccttattg acatttgagg ggctgtccac aggcagaaaa tccagcattt gcaagggttt 2820
 cgcgcgttt ttccggccacc gctaacctgt cttaaacctt gcttttaaac caatatttat 2880
 aaaccttgtt tttaaccagg gctgcgcacct gtgcgcgtga cgcgcacgc cgaaggggg 2940
 tgccccccct ttcgaaccc tcccgccccg ctaacgcggg ctcgcaccc cccaggggc 3000
 tgcgcgcctc ggcgcgaac ggcctcaccc caaaaatggc agcgctggca gtccctgcca 3060
 ttgcgggat cggggcagta acgggatggg cgatcagccc gagcgcgacg cccggaagca 3120
 ttgacgtgcc gcaggtgctg gcatcgacat tcagcgacca ggtgcgggc agtgagggcg 3180
 gggcctggg tggcggcctg cccttcaattt cggccgtcgg ggcattcacg gacttcatgg 3240
 cggggccggc aatttttacc ttggcattt tcggcatagt ggtgcgggt gccgtgctcg 3300
 ttttcgggg tgcgataaac ccagcgaacc attttaggtg ataggtaaga ttataccgag 3360
 gtatgaaaac gagaattgga ctttacaga attactctat gaagcgccat attaaaaaag 3420
 ctaccaagac gaagaggatg aagaggatga ggaggcagat tgccttgaat atattgacaa 3480
 tactgataag ataatatatac ttttatatac aagatatcgc cgtatgtaa gatttcaggg 3540
 ggcaaggcat aggcagcgcg cttatcaata tatctataga atggcaaag cataaaaaact 3600
 tgcattggact aatgcattgaa acccaggaca ataacctt atgcttgcattt ttctatcata 3660
 attgggtaat gactccaact tattgatagt gttttatgtt cagataatgc ccgatgactt 3720
 tgtcatgcag ctccaccgat ttggagaacg acagcgactt ccgtcccagc cgtgccaggt 3780
 gctgcctcag attcaggtta tgccgcctaa ttcgcgtatgcgat atatcgcttgcgt 3840
 gcagcttcc ctccaggcgg gattcataca gcccgcggcc atccgtcattc catatcacca 3900
 cgtcaaaaggg tgacagcagg ctcataagac gccccagcgt cgccatagtg cggttcaccga 3960
 atacgtgcgc aacaaccgtc ttccggagac tgtcatacgc gtaaaacagc cagcgctggc 4020
 gcgatttagc cccgacatag ccccaactgtt cgtccatttc cgccgcacgc atgacgtcac 4080
 tgcccggttg tatgcgcgag gttaccgact gccgcgtgag ttttttaagt gacgtaaaat 4140

cgttgtgagg ccaacgccc taatgcggc tggcccg catccaacgc cattcatggc 4200
 catatcaatg atttctggt gcgtaccggg ttgagaagcg gtgttaagtga actgcagttg 4260
 ccatgtttta cggcagttag agcagagata gcgcgtatgt ccggcggtgc tttgccgtt 4320
 acgcaccacc ccgtcagtag ctgaacagga gggacagctg atagacacag aagccactgg 4380
 agcacctcaa aaacaccatc atacactaaa tcagtaagtt ggcagcatca cccataattg 4440
 tggtttcaaa atcggtccg tcgatactat gttatacgcc aactttgaaa acaactttga 4500
 aaaagctgtt ttctggtatt taaggtttta gaatgcaagg aacagtgaat tggagttcgt 4560
 cttgttataa ttagttctt ggggtatctt taaatactgt agaaaaagagg aaggaaataa 4620
 taaatggcta aaatgagaat atcaccggaa ttgaaaaaac tgatcgaaaa ataccgctgc 4680
 gtaaaagata cggaaggaat gtctccgt aaggtatata agctggggg agaaaaatgaa 4740
 aacctatatt taaaaatgac ggacagccgg tataaaggga ccacctatga tgtggaacgg 4800
 gaaaaggaca tgatgctatg gctggaagga aagctgcctg ttccaaaggt cctgcacttt 4860
 gaacggcatg atggctggag caatctgctc atgagtgagg ccgatggcgt ccttgctcg 4920
 gaagagtatg aagatgaaca aagccctgaa aagattatcg agctgtatgc ggagtgcac 4980
 aggcttttc actccatcga catatcgat tgcctata cgaatagctt agacagccgc 5040
 ttagccaat tggattactt actgaataac gatctggccg atgtggattg cgaaaactgg 5100
 gaagaagaca ctccatttaa agatccgcgc gagctgtatg atttttaaa gacggaaaag 5160
 cccgaagagg aacttgcctt ttcccacggc gacctgggag acagcaacat cttgtgaaa 5220
 gatggcaaag taagtggctt tattgatctt gggagaagcg gcagggcggc caagtggat 5280
 gacattgcct tctgcgtccg gtcgatcagg gaggatatcg gggagaaca gtaatgtcgag 5340
 ctatttttg acttactgg gatcaagcct gattgggaga aaataaaata ttatattta 5400
 ctggatgaat tggtagta cctagatgtg ggcacacgt gcccgcgaca agcaggagcg 5460
 caccgacttc ttccgcac 1ca agtgtttgg ctctcaggcc gaggcccacg gcaagtattt 5520
 gggcaagggg tcgtgttat tcgtgcaggg caagattcgg aataccaagt acgagaagga 5580
 cggccagacg gtctacggga ccgacttcat tgccgataag gtggattatc tggacaccaa 5640
 ggcaccaggg gggtaatgc atcggacgtt tgaccggaag gcatacaggc aagaactgtat 5700
 cccgcaaggg gggtaatgc atcggacgtt tgaccggaag gcatacaggc aagaactgtat 5760
 cgacgcgggg tttccgcgg aggtgccga aaccatcgca agccgcaccc tcacgcgtgc 5820
 gccccgcgaa accttccagt ccgtcggctc gatggtccag caagctacgg ccaagatcg 5880

gcgcgacagc gtgcaactgg ctccccctgc cctgcccgcg ccatcgcccg ccgtggagcg 5940
ttcgcgtcgt ctcgaacagg aggccggcagg tttggcgaag tcgatgacca tcgacacgcg 6000
aggaactatg acgacccaaga agcgaaaaac cgccggcggag gacctggcaa aacaggtcag 6060
cgaggccaag caggccgcgt tgctgaaaca cacgaagcag cagatcaagg aaatgcagct 6120
ttccttgttc gatattgcgc cgtggccgga cacgatgcga gcgatgcca aacgacacggc 6180
ccgctctgcc ctgttcacca cgcgcaacaa gaaaatcccg cgcgaggcgc tgcaaaacaa 6240
ggtcattttc cacgtcaaca aggacgtgaa gatcacctac accggcgtcg agctgcggc 6300
cgacgatgac gaactggtgt ggcagcaggt gttggagtac gcgaagcgc a cccctatcg 6360
cgagccgatc accttcacgt tctacgagct ttgccaggac ctgggctggc cgatcaatgg 6420
ccggtattac acgaaggccg aggaatgcct gtcgcgccta cagggcgcacgg cgatggcctt 6480
cacgtccgac cgcgttggc acctggaatc ggtgtcgctg ctgcaccgc tccgcgtcct 6540
ggaccgtggc aagaaaaacgt cccgttgcga ggtcctgatc gacgagaaaa tcgtcgtgct 6600
gtttgctggc gaccactaca cgaaattcat atgggagaag taccgcaagc tgtcgccgac 6660
ggcccgacgg atgttcgact atttcagctc gcacccggag ccgtacccgc tcaagctgga 6720
aaccttccgc ctcatgtgcg gatcgattc cacccgcgtg aagaagtggc gcgagcaggt 6780
cggcgaagcc tgcgaagagt tgcgaggcag cggcctggc gaacacgcct gggtaatga 6840
tgacctggc cattgcaaac gctagggcct tgcgggtca gttccggctg ggggttcagc 6900
agccagcgc ttaactggcat ttccaggaaca agcgggcact gtcgacgc cttgcgtcgc 6960
tcagtatcgc tcgggacgc cggcgcgc taegaactgc cgataaacag aggattaaaa 7020
ttgacaattt tgattaaggc tcagattcga cggcgttggag cggccgcacgt gcaggatttc 7080
cgcgagatcc gattgtcgcc cctgaagaaa gctccagaga tgcgggtca cgatggcattc 7140
cacgaggaga aaaagcccat ggaggcgttc gctgaacggc tgcgagatgc cgtggcattc 7200
ggcgccata tcgacggcga gatcattggg ctgtcggtct tcaaacagga ggacggcccc 7260
aaggacgc tcaaggcgc tctgtccggc gtttgcgtgg agccgcacaa ggcggccg 7320
ggggtcgccc gtatgcgtct gccccgggtt ccggccgggtt tattgcgtt gatgtcg 7380
cgacagattc caacggaaat ctgggtggatc cgcacatctca tcctcgccgc acttaatatt 7440
tcgctattct ggagcttgc ttttgcgtt gtcgtaccgc tgccggccgg ggtcgccgc 7500
acggtaggcg ctgtgcagcc gctgtggcgt gtgttcatct ctgcgcgtct gctaggtac 7560
ccgatacgt tgatggcggt cctggggct atttgcggaa ctgcggcggt ggcgtgttgc 7620

gtgttacac caaacgcagc gctagatcct gtcggcgctc cagcggccct ggccccggc 7680
gtttccatgg cggtcgaaac cgtgctgacc cgcaagtggc aaccccccgt gcctctgctc 7740
acctttacccg cctggcaact ggccggccggc ggacttctgc tggcccttggc agcttttagtg 7800
tttgatccgc caatcccgat gcctacagga accaatgttc tggccctggc gtggctcgcc 7860
ctgatcgag cgggttaac ctacttcctt tggccctggg ggatctcgcc actcgaacct 7920
acagttgtt ccttactggg ctttctcaggc cccagatctg gggtcgatca gccggggatg 7980
catcaggccg acagtcggaa ctgcgggtcc ccgacctgtt ccattcggtg agcaatggat 8040
aggggagttg atatcgtaa cgttcacttc taaagaaaata gcccactca gcttcctcag 8100
cggtttatc cagcgatttc ctattatgtc ggcatacgatc tcaagatcga cagcctgtca 8160
cggttaagcg agaaatgaat aagaaggctg ataattcggg tctctgcgag ggagatgata 8220
tttgatcaca ggcagcaacg ctctgtcatc gttacaatca acatgctacc ctccgcgaga 8280
tcatccgtgt ttcaaaccggc gcaagcttagt tgccgttctt ccgaatagca tcggtaacat 8340
gagcaaagtc tgccgcctta caacggctct cccgctgacg ccgtcccgga ctgatgggct 8400
gcctgtatcg agtggtgatt ttgtgccgag ctgcgggtcg gggagctgtt ggctggctgg 8460
tggcaggata tattgtggtg taaacaaatt gacgcttaga caacttaata acacattgcg 8520
gacgaaaaatc atgtactggg gtggtttttc ttttaccagg tgagacgggc aacagctgat 8580
tgcccttcac cgcctggccc tgagagagtt gcagcaagcg gtccacgctg gttggccca 8640
gcagggcgaaa atccgttttgc atgggtggttc cgaaatcgcc aaaatccctt ataaatcaaa 8700
agaatagccc gagataggggt tgagtgttgt tccagttgg aacaagagtc cactattaaa 8760
gaacgtggac tccaacgtca aagggcgaaa aaccgtctat cagggcgatg gcccactacg 8820
tgaaccatca cccaaatcaa gtttttggg gtcgagggtgc cgtaaagcac taaatcgaa 8880
ccctaaaggg agcccccgat ttagagcttgc acggggaaag ccggcgaacg tggcgagaaa 8940
ggaagggaag aaagcgaaag gagcggcgcc cattcaggct ggcacactgt tgggaaggc 9000
gatcggtgcg ggcccttcgc ctattacgccc agctggcgaa agggggatgt gctgcaaggc 9060
gattaagttg ggtAACGCCA gggtttccc agtcacgacg ttgtaaaacg acggccagtg 9120
aattaattcc catottgaaa gaaatatagt ttaaatattt attgataaaaa taacaagtca 9180
ggtattatag tccaagcaaa aacataaatt tattgatgca agttaaatt cagaaatatt 9240
tcaataactg attatatcag ctggtagatc gccgttagatg aaagactgag tgcgatatta 9300
tgtgtataac ataaatttgc gatatagctca gcttagctca tcgggggatc cgtcgaagct 9360

agcttgggtc cogctcgaaa gaactcgta agaaggcgat agaaggcgat ggcgctcgaa 9420
tcgggagccg cgataaccgta aagcacgagg aagcggtcag cccattcgcc gccaagctct 9480
tcagcaatat cacgggttagc caacgctatg tcctgatagc ggtccgccac acccagccgg 9540
ccacagtcga tgaatccaga aaagcgccca ttttccacca tgatattcgg caagcaggca 9600
tcgcatggg tcacgacgag atcctcgccg tcgggcatgc ggcgccttgag cctggcgaac 9660
atttcggctg ggcgcagccc ctgatgtct tcgtccagat catcctgatc gacaagacgg 9720
gcttccatcc gagtacgtgc tagtcgtatc cgatgtttcg cttgggtggtc gaatgggcag 9780
gtagccggat caagcgatcg cagccgccc attgcattcg ccatgtatgg tactttctcg 9840
gcaggagcaa ggtgagatga caggagatcc tgccccggca cttcgcccaa tagcagccag 9900
tccctcccg cttcagtgac aacgtcgagc acagctgcgc aaggaacgcc cgtcgtggcc 9960
agccacgata gccgcgctgc ctgcgtctgc agttcattca gggcaccggc caggtcggtc 10020
ttgacaaaaaa gaaccgggcg cccctgcgtc gacagccggc acacggccgc atcagagcag 10080
ccgattgtct gttgtccca gtcatagccg aatagcctct ccacccaagc ggccggagaa 10140
cctgcgtgca atccatcttg ttcaatccaa gctccatgg gccctcgact agagtcgaga 10200
tctggattga gagtgaatat gagactctaa ttggataccg agggaaattt atggaacgtc 10260
agtggagcat ttttgacaag aaatatttgc tagctgatag tgaccttagg cgactttga 10320
acgcgcata atggttctg acgtatgtgc ttagctcatt aaactccaga aaccgcggc 10380
tgagttggctc cttcaacgtt gcggttctgt cagttccaaa cgtaaaacgg cttgtccgc 10440
gtcatcgccg ggggtcataa cgtgactccc ttaattctcc gctcatgatc ttgatcccct 10500
gcgcctcag atccctggcg gcaagaaaagc catccagttt actttgcagg gctcccaac 10560
cttaccagag ggcccccag ctggcaattc cggttcgctt gctgtccata aaaccgcggc 10620
gtctagctat cgccatgtaa gcccactgca agtacactgc ttctctttg cgcttgcgtt 10680
ttcccttgc cagatagccc agtagctgac attcatccgg ggtcagcacc gtttctgggg 10740
actggcttc taegtgttcc gttccattt gcagcccttg cgccctgagt gcttgccggc 10800
gcgtgaagct tgcatgcctg caggtcgacg ggcgcggcag ctccctcgagc aaatttacac 10860
attgccacta aacgtctaaa cccttgcata ttgttttgc ttactatgt gtgttatgt 10920
tttgatggc gataaatttt tatatttggt actaaattta taacaccctt tatgctaacg 10980
tttgccaaaca ctttagcaatt tgcaagttga ttaatttgcatt ctaaatttattttt tttgtcttct 11040
aaatacatat actaatcaac tggaaatgt aatatttgc aatatttcta ctataggaga 11100

attaaagtga gtgaardatgg taccacaagg tttggagatt taattgttgc aatgctgcat 11160
 ggtatggcata tacaccaaac attcaataat tcttgaggat aataatggta ccacacaaga 11220
 tttgaggtgc atgaacgtca cgtggacaaa aggttttagta attttcaag acaacaatgt 11280
 taccacacac aagtttgag gtgcattgcat ggatgccctg tggaaagttt aaaaatattt 11340
 tggaaatgat ttgcattgaa gccatgtgta aaaccatgac atccacttgg aggatgcaat 11400
 aatgaagaaa actacaaatt tacatgcaac tagttatgca tgttgtctat ataatgagga 11460
 ttttgcata ctttcattca tacacactca ctaagttta cacgattata atttcttcatt 11520
 agccagcgga tccgatatcg ggcccgtag cgtaaccct gcttaatga gatatgcgag 11580
 acgcctatga tcgcattgata tttgcattca attctgttgc gcacgttgc aaaaacctga 11640
 gcatgtgttag ctcagatcct taccgcggc ttccgttcat tctaattaatgat atatcacccg 11700
 ttactatcgat atttttatgat ataataatttccgatt tactgattgt ccgtcgacga 11760
 attcgagctc ggcgcgcctc tagaggatcg atgaattcag atcggctgag tggctccctc 11820
 aacgttgcgg ttctgtcagt tccaaacgta aaacggcttgc tccgcgtca tcggcggggg 11880
 tcataaacgtg actcccttaa ttctccgctc atgatcagat tgctgttcc cgccttcagt 11940
 ttaaactatc agtgtttgac aggatatttggcggtaaa cctaagagaa aagagcggtt 12000
 attagaataa tcggatattt aaaaggcgt gaaaagggtt atccttcgtc catttgtatg 12060
 tgcatgccaa ccacagggtt cccca 12085

<210> 22
 <211> 12079
 <212> DNA
 <213> Unknown

<220>

<223> pflanzlicher Expressionsvektor mit einer Promotor-Terminator-Expressionskassette

<400> 22

gatctggcgc cggccagcga gacgagcaag attggccgccc gccccaaacg atccgacagc 60
 gcccggcagca caggtgcgcga ggcaattgc accaacgcattt acagcgccag cagaatgc 120
 tagtggccgg tgacgtcggtt cgagtgaacc agatcgccga ggaggccgg cagcaccggc 180
 ataatcaggc cgtatccgcac agcgtcgagc ggcacagtgc tcagaattac gatcagggtt 240
 atgttgggtt tcacgtctgg cctccggacc agcctccgct ggtccgattt aacgcgcgg 300
 ttctttatca ctgataagtt ggtggacata ttatgtttat cagtgataaaa gtgtcaagca 360
 tgacaaagtt gcagccgaat acagtgtatcc gtgcgcctt ggacctgttgc aacgaggcgtc 420

gcgttagacgg tctgacgaca cgcaaactgg cgaaacggtt gggggttcag cagccggcgc 480
 ttactggca cttcaggaac aagegggcgc tgctcgacgc actggccgaa gccatgctgg 540
 cggagaatca tacgcattcg gtgccgagag ccgacgacga ctggcgctca tttctgatcg 600
 ggaatcccc cagcttcagg cagggcgctgc tggcctaccg cgatggcgcg cgcatccatg 660
 cggcacgcg accggggcgca ccgcagatgg aaacggccga cgcgagctt cgcttcctct 720
 gcgaggcggg ttttcggcc ggggacgccc tcaatgcgct gatgacaatc agctacttca 780
 ctgttggggc cgtgcttgag gagcaggccg gcgcacagcga tgccggcgag cgccggggca 840
 ccgttgaaca ggctccgctc tcgcgcgtgt tgccggccgc gatagacgcc ttgcacgaag 900
 ccggtccgga cgcagcgttc gagcagggac tcgcggtgat tgtcgatgga ttggcgaaaa 960
 ggaggcttgt tgtcaggaac gttgaaggac cgagaaaaggg tgacgattga tcaggacccgc 1020
 tgccggagcg caacccactc actacagcag agccatgtag acaacatccc ctcccccttt 1080
 ccaccgcgtc agacgcccgt agcagccgc tacggcttt ttcatgccc gcctagcgt 1140
 ccaaggctca cggccgcgct cggcctctct ggcggccctc tggcgctctt ccgttccctc 1200
 gctactgac togctgcgct cggtcgttcg gctgcggcga gcggtatcag ctcaactaaa 1260
 ggccgtaata cggttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa 1320
 aggccagcaa aaggccagga accgtaaaaa gcccgcgttg ctggcggttt tccataggct 1380
 ccgccccctt gacgagcatc acaaaaatcg acgctcaagt cagagggtggc gaaaccgcac 1440
 aggactataa agataaccagg cgtttccccc tggaaagctcc ctgcgtgcgt ctctgttcc 1500
 gaccctgccc cttaccggat acctgtccgc ctttctccct tcgggaagcg tggcgctttt 1560
 cccgtgcata accctgcttc ggggtcatta tagcgatttt ttcggtatata ccattctttt 1620
 tcgcacgata tacaggattt tgccaaaggg ttctgtttaga ctttccttgg tgtatccaac 1680
 ggcgtcagcc gggcaggata ggtgaagtag gcccacccgc gagcgggtgt tccattttca 1740
 ctgtccctta ttccgcacctg gcggtgctca acgggaatcc tgctctgcga ggctggccgg 1800
 ctaccgcggc cgtaacagat gagggaacgc ggatggctga tgaaaccaag ccaaccagga 1860
 agggcagccc acctatcaag gtgtactgcc ttccagacga acgaagagcg attgaggaaaa 1920
 aggcggcggc ggccggcatg agcctgtcgg cctacctgct ggccgtcggc cagggtaca 1980
 aaatcacggg cgtcgtggac tatgagcagc tccgcgagct ggcccgcatc aatggcgacc 2040
 tgggcgcctt gggcgccctg ctgaaaactct ggctcaccga cgacccgcgc acggcgccgt 2100
 tcggtgatgc cacgatcctc gcccgtctgg cgaagatcga agagaagcag gacgagcttgc 2160

gcaaggcat gatggcggtg gtccgcgcga gggcagagcc atgactttt tagccgctaa 2220
aacggccggg gggtgcgcggt gattgccaag cacgtccccat tgcgctccat caagaagagc 2280
gacttcgcgg agctggtaa gtacatcacc gacgagcaag gcaagaccga gcgccttgc 2340
gacgctcacc gggctggttt ccctcgccgc tggctggcg gccgtctatg gccctgaaa 2400
cgcgccagaa acgcccgtcga agccgtgtgc gagacaccgc ggccgcggc gttgtggata 2460
cctcgccgaa aacttggccc tcactgacag atgaggggcg gacgttgaca cttgagggc 2520
cgactcaccc ggcgcggcggt tgacagatga ggggcaggct cgatttcggc cggcgacgtg 2580
gagctggcca gcctcgcaaa tcggcgaaaa cgccctgatt tacgctgagtt tcccacagat 2640
gatgtggaca agcctgggaa taagtgcctt ggggtattga cacttgaggg gcgactac 2700
tgacagatga ggggcgcgat ctttgacact tgagggcag agtgcgtaca gatgagggc 2760
gcacctattt acatttgagg ggctgtccac aggcaaaaa tccagcattt gcaagggttt 2820
ccgccccgttt ttccggccacc gctaacctgt cttaacactt gcttttaaac caatatttat 2880
aaaccttgtt tttaaccagg gctgcgcctt gtgcgcgtga ccgcgcacgc cgaagggggg 2940
tgccccccct tctcgaaaccc tccccggcccg ctaacgcggg cctccatcc cccagggc 3000
tgcccccctc ggccgcgaac ggctcaccc caaaaatggc agcgctggca gtccttgcca 3060
ttgccgggat cggggcagta acgggatggg cgatcagccc gagcgcgacg cccgaaagca 3120
ttgacgtgcc gcaggtgctg gcatcgacat tcagcgacca ggtgcgggc agtggggcg 3180
gcggccctggg tggccgcctg ccttcactt cggccgtcg ggcattcacg gacattcatgg 3240
cgggccggc aatttttacc ttggcattt ttggcatagt ggtgcgggt gccgtgctcg 3300
tgttcggggg tgcgataaac ccagcgaaacc atttgagggtg ataggtaaga ttataccgag 3360
gtatgaaaac gagaatttggaa cttttacaga attactctat gaagcgccat attaaaaag 3420
ctaccaagac gaagaggatg aagaggatga ggaggcagat tgccttgaat atattgacaa 3480
tactgataag ataatatatc ttttatatag aagatatcgc cgtatgtaa gatttcaggg 3540
ggcaaggcat aggcagcgcg cttatcaata tatctataga atgggcaaag cataaaaaact 3600
tgcatggact aatgcttggaa acccaggaca ataaccttat agcttggaaa ttctatcata 3660
attgggtaat gactccaact tattgatagt gtttatgtt cagataatgc ccgtatgactt 3720
tgtcatgcag ctccaccgat tttgagaacg acagcgactt ccgtcccagc cgtgccaggt 3780
gtgcctcag attcaggta tgccgctcaa ttgcgtcggt atatcgctt ctgattacgt 3840
gcagcttcc ctccaggcg gattcataca gcggccagcc atccgtcatac catatcacca 3900

cgtcaaaggg tgacagcagg ctcataagac gccccagcgt cgccatagtgcg 3960
 atacgtgcgc aacaaccgtc ttccggagac tgtcatacgc gtaaaacagc cagcgctggc 4020
 gcgatttagc cccgacatag ccccactgtt cgtccatttc cgccgacagc atgacgtcac 4080
 tgccggctg tatgcgcgag gttaccgact gcggccttag ttttttaagt gacgtaaaat 4140
 cgtgttgagg ccaacgccc taatgcgggc tggcccccgc catccaacgc cattcatggc 4200
 catatcaatg attttctggt gcgtaccggg ttgagaagcg gtgtaaatgtactgcgttgc 4260
 ccatgtttta cggcagttag agcagagata gcgtgtatgt ccggcggtgc ttttgcgtt 4320
 acgcaccacc cggcgttag ctgaacagga gggacagctg atagacacag aagccactgg 4380
 agcacctcaa aaacaccatc atacactaaa tcagtaagtt ggcagcatca cccataattg 4440
 tggtttcaaa atcggctccg tcgataactat gtatacgcc aactttgaaa acaactttga 4500
 aaaagctgtt ttctggtatt taaggtttta gaatgcagg aacagtgaat tggagttcgt 4560
 ctgttataa tttagttctt ggggtatctt taaatactgt agaaaagagg aaggaaataa 4620
 taaatggcta aaatgagaat atcaccggaa ttgaaaaaac tgatcgaaaa ataccgctgc 4680
 gtaaaaagata cggaaaggaat gtctccgt aaggtatata agctggggg agaaaatgaa 4740
 aacctatatt taaaaatgac ggacagccgg tataaaggga ccacctatga tgtggAACGG 4800
 gaaaaggaca tgatgtatgc gctggagg aagctgcctg ttccaaaggt cctgcacttt 4860
 gaacggcatg atggctggag caatctgctc atgagtggagg ccgtggcgt cctttgctcg 4920
 gaagagtatg aagatgaaca aagccctgaa aagattatcg agctgtatgc ggagtgcac 4980
 aggcttttc actccatcgat catatcgat tgtccctata cgaatagctt agacagccgc 5040
 ttagccaat tggattactt actgaataac gatctggccg atgtggatttgcgaaaactgg 5100
 gaagaagaca ctccattnaa agatccgcgc gagctgtatg atttttaaa gacggaaaag 5160
 cccgaagagg aacttgtctt ttcccacggc gacctggggag acagcaacat ctttgtaaa 5220
 gatggcaaag taagtggctt tattgatctt gggagaagcg gcagggcggca aagtggtat 5280
 gacattgcct tctgcgtccg gtcgatcagg gaggatatcg gggagaaca gtatgtcgag 5340
 ctatTTTGTG acttactggg gatcaaggct gattgggaga aaataaaaata ttatattta 5400
 ctggatgaat tgttttagta cctagatgtg ggcgcacat gcccggcaca agcaggagcg 5460
 caccgacttc ttccgcata gatgtttgg ctctcaggcc gaggcccacg gcaagtattt 5520
 gggcaagggg tcgctggat tcgtgcaggca aatccaatgt acgagaagga 5580
 cggccagacg gtctacggga ccgacttcat tgccgataag gtggattatc tggacaccaa 5640

ggcaccaggc gggtaaaatc aggaataagg gcacattgcc ccggcgtag tcggggcaat 5700
 cccgcaagga gggtaatga atcgacgtt tgaccgaaag gcatacaggc aagaactgat 5760
 cgacgcgggg tttccgccc aggatgccga aaccatcgca agccgcaccc tcatgcgtgc 5820
 gccccgcgaa accttccagt ccgtcggctc gatggtccag caagctacgg ccaagatcga 5880
 gcgcgacagc gtgcaactgg ctccccctgc cctgcccgcg ccatcggccg ccgtggagcg 5940
 ttgcgcgtcg ttcgaacagg aggccgcagg tttggcgaag tcgatgacca tcgacacgcg 6000
 aggaactatg acgacccaaga agcgaaaaac cggccgcgag gacctggcaa aacaggtcag 6060
 cgaggccaag caggccgcgt tgctgaaaca cacgaagcag cagatcaagg aaatgcagct 6120
 ttcctgttc gatattgcgc cgtggccgga cacgatgcga gcatgcacca acgacacggc 6180
 ccgcctgcg ctgttcacca cgcgcaacaa gaaaatccc cgcgaggcgc tgcaaaacaa 6240
 ggtcatttc cacgtcaaca aggacgtgaa gatcacctac accggcgctcg agctgcgggc 6300
 cgacgatgac gaactggtgt ggcagcaggt gttggagtac gcatgcgcgca cccctatcg 6360
 cgagccgatc accttcacgt tctacgagct ttgccaggac ctgggcttgt cgatcaatgg 6420
 ccggtattac acgaaggccg aggaatgcct gtcgcgccta caggcgacgg cgatggctt 6480
 cacgtccgac cgcgttgggc acctggaatc ggtgtcgctg ctgcaccgct tccgcgtcct 6540
 ggaccgtggc aagaaaaacgt cccgttgcga ggtcctgatc gacgaggaaa tcgtcgctg 6600
 gtttgcgtgc gaccactaca cgaattatcat atgggagaag taccgcaagc tgtcgcgcac 6660
 ggccccacgg atgttcgact atttcagctc gcaccggag ccgtaccgcg tcaagctgga 6720
 aaccttccgc ctcatgtcg gatcgattc caccgcgtg aagaagtggc gcgagcaggt 6780
 cggcgaagcc tgcgaaagagt tgcgaggcag cggcctggtg gaacacgcct gggtaatga 6840
 tgacctggtg cattgcaaac gctagggcct tgtgggtca gttccggctg ggggttcagc 6900
 agccagcgct ttactggcat ttcaagaaaca agcgggact gtcgacgcg cttgcttcgc 6960
 tcagtatcgc tcgggacgcg cggcgcgcgc tacgaactgc cgataaacag aggattaaaa 7020
 ttgacaattg tgattaaggc tcagattcga cggctggag cggccgacgt gcaggatttc 7080
 cgcgagatcc gattgtcgcc cctgaagaaa gctccagaga tgttcgggtc cgtttacgag 7140
 cacgaggaga aaaagcccat ggaggcggtc gctgaacggc tgcgagatgc cgtggcatc 7200
 ggcgcctaca tcgacggcgatc gatcattggg ctgtcggtct tcaaacagga ggacggcccc 7260
 aaggacgctc acaaggcgatc tctgtccggc gtttcgtgg agcccgaaaca gcgaggccgatc 7320
 ggggtcgccg gatatgtcgatc gggggcggtt tattgctcgatc gatgtcgatc 7380

cgacagattc caacgggaat ctggtgatg cgcatcttca tccctggcgc acttaatatt 7440
 tcgctattct ggagcttgtt gtttatttcg gtctaccgcc tgccggcg 7500
 acggtaggcg ctgtgcagcc gctgatggtc gtgttcatct ctgcccgtct gctaggtac 7560
 ccgatacgt tggatggcggt cctggggct atttgcggaa ctgccccgtt ggcgcgtttg 7620
 gtgttgcacac caaacgcagc gctagatcct gtcggcgac cagccccctt ggccccggcg 7680
 gtttccatgg cggtcgaaac cgtgctgacc cgcaagtggc aacccccgtt gcctctgttc 7740
 acctttaccg cctggcaact ggcggccgga ggacttctgc tggatccagt agcttttagtg 7800
 tttgatccgc caatcccgat gcctacagga accaatgttc tggccctggc gtggctcg 7860
 ctgatcgag cgggtttaac ctacttcctt tggttccggg ggatctcg 7920
 acagttgttt ctttactggg ctttctcagc cccagatctg ggtcgatca gccggggatg 7980
 catcaggccg acagtcggaa ctccgggtcc ccgacctgtt ccattcggtt agcaatggat 8040
 aggggagttt atatcgtaa cggttacttc taaagaaata ggcactca gtttcctcag 8100
 cggctttatc cagcgatttc ctattatgtc ggcatagttc tcaagatcga cagcctgtca 8160
 cggtaagcg agaaatgaat aagaaggctt ataattcgga tctctgcgag ggagatgata 8220
 tttgatcaca ggcagcaacg ctctgtcatc gttacaatca acatgctacc ctccgcgaga 8280
 tcattccgtt ttcacaaacccg gcagcttagt tgccgttctt ccgaatagca tggtaacat 8340
 gagcaaagtc tgccgccta caacggctct cccgctgacg ccgtccccgaa ctgatggct 8400
 gcctgtatcg agtggtgatt ttgtggcgag ctggcggtcg gggagctgtt ggctggctgg 8460
 tggcaggata tattgtggtg taaacaaatt gacgcttaga caacttaata acacattgcg 8520
 gacgtttta atgtactggg gtggttttc ttttccaccag tgagacgggc aacagctgat 8580
 tgcccttca cgcctggccc tgagagagtt gcagcaagcg gtccacgctg gtttgcggca 8640
 gcaggcgaaa atcctgtttt atgggggttc cggaaatcggtt aaaatccctt ataaatcaaa 8700
 agaatacgccc gagatagggt tgagtgttgc tccagtttgg aacaagagtc cactattaaa 8760
 gaacgtggac tccaacgtca aaggcgaaa aaccgtctat cagggcgatg gcccactacg 8820
 tgaaccatca cccaaatcaa gtttttggg gtggcggtgc cgtaaagcac taaatcgaa 8880
 ccctaaaggag gccccccgtt ttagagcttgc acggggaaag ccggcgaaacg tggcgagaaa 8940
 ggaaggaaag aaagcgaaaag gagcggcgcc cattcaggct ggcactgtt tggtggggc 9000
 gatcggtgcg ggccttttcg ctattacgccc agctggcgaa agggggatgt gctgcaaggc 9060
 gattaagttt ggttaacgcca gggttttccc agtcacgacg ttgtaaaacg acggccagtg 9120

aattaattcc catcttgaaa gaaatatagt ttaaatatTTT attgataaaaa taacaagtca 9180
 ggtattatag tccaagcaaa aacataaatt tattgatgca agtttaaatt cagaatatt 9240
 tcaataactg attatATCAG ctggtagatt gccgtatgt aaagactgag tgcatatta 9300
 tgtgtatac ataaattgtat gatatagcta gcttagctca tcggggatc cgtcgaagct 9360
 agcttgggtc ccgctcagaa gaactcgta agaaggcgat agaaggcgat gcgcgcgaa 9420
 tcgggagcg cgataaccgta aagcacgagg aagcggtcag cccattcgcc gccaagctct 9480
 tcagcaatat cacgggttagc caacgctatg tcctgatagc ggtccgccc acccagccgg 9540
 ccacagtcga tgaatccaga aaagcgccca tttccacca tgatattcgg caagcaggca 9600
 tcgccccatggg tcacgacgag atcctcgccg tcggcatgc gcgcctttag cctggcgaac 9660
 agttcggctg ggcgcgagccc ctgatgtct tcgtccagat catcctgatc gacaagaccg 9720
 gcttccatcc gagtacgtgc tcgctcgatg cgatgtttcg cttgggtggc gaatggcag 9780
 gtagccggat caagcgatcg cagccgcgc attgcatcag ccatgatgga tactttctcg 9840
 gcaggagcaa ggtgagatga caggagatcc tgccccggca cttcgcggaa tagcagccag 9900
 tccctcccg cttcagtgc acacgtcgac acagctcgcc aaggaacgcc .cgtcgtggcc 9960
 agccacgata gcccgcgtgc ctgttcctgc agttcattca gggcaccggc caggtcggtc 10020
 ttgacaaaaaa gaaccggggcg cccctgcgt gacagccggc acacggccggc atcagagcag 10080
 ccgattgtct gttgtgcccc gtcatagccg aatagcctct ccacccaaagc ggccggagaa 10140
 cctgcgtgca atccatcttgc ttcaatccaa gctccatgg gcccctcgact agagtcgaga 10200
 tctggattga gagtgaatat gagactctaa ttggataccg aggggaattt atggaacgtc 10260
 agtggagcat ttttgacaag aaatatttgc tagctgatag tgacctttagg cgactttga 10320
 acgcgcataa atggtttctg acgtatgtgc tttagctcatt aaactccaga aaccggccggc 10380
 tgagtggctc cttcaacgtt gcggtctgt cagttccaaa cgtaaaaacgg cttgtcccg 10440
 gtcatcgccg ggggtcataa cgtgactccc ttaattctcc gctcatgatc ttgatcccc 10500
 ggcgcatac atccttggcg gcaagaaagc catccagttt actttgcagg gctcccaac 10560
 cttaccagag ggcgcggccag ctggcaattc cggttcgtt gctgtccata aaaccggcc 10620
 gtcttagctat cgccatgtaa gcccactgca agtacactgc tttcttttgc cgcttgcgtt 10680
 ttcccttgc cagatagccc agtagctgac attcatccgg ggtcagcacc gtttctgcgg 10740
 actggcttc tacgtgttcc gtttcatttgc agggcccttgc gcccctgagt gctgcggca 10800
 gcgtgaagct tgcatgcctg caggtcgacg ggcgcggccag ctcctcgagc aaatttacac 10860

attgccacta aacgtctaaa cccttgcata ttgttttgt tttactatgt gtgttatgta 10920
 tttgatttgc gataaaatttt tatatttggt actaaattta taacaccctt tatgctaacg 10980
 tttgccaaca cttagcaatt tgcaagttga ttaattgatt ctaaatttatt tttgtcttct 11040
 aaatacatat actaatcaac tggaaatgt aatatttgc aatatttcta ctataggaga 11100
 attaaagtga gtgaatatgg taccacaagg tttggagatt taattgtgc aatgctgcat 11160
 ggtatggcata tacacccaaac attcaataat tcttgggat aataatggta ccacacaaga 11220
 tttgaggtgc atgaacgtca cgtggacaaa agtttagta attttcaag acaacaatgt 11280
 taccacacac aagtttttagt gtgcattgc ggtatgcctg tggaaagttt aaaaatattt 11340
 tggaaatgat ttgcatggaa gccatgtgt aaccatgac atccacttgg aggatgcaat 11400
 aatgaagaaa actacaaatt tacatgcaac tagttatgca tgtgtctat ataatgagga 11460
 ttttgcata ctttcattca tacacactca ctaagttta cacgattata atttcttcat 11520
 agccagcaga tctgcccggca tcgatccccgg gcatggcct gcttaatgt gatatgcgag 11580
 acgcctatga tcgcatgata tttgcttca attctgttgt gcacgttgc aaaaacctga 11640
 gcatgtgttag ctcagatccc taccggccgt ttcggttcat tctaattgtat atatcaccgg 11700
 ttactatcgat atttttatgtataatattct ccgttcaatt tactgattgt ccgtcgacga 11760
 gctcgccgcg cctctagagg atcgatgtat tcagatccgc tgagtggctc cttcaacgtt 11820
 gcggttctgt cagttccaaa cgtaaaaacgg ctgtcccg cgtatccgc ggggtcataa 11880
 cgtgactccc ttaattctcc gctcatgatc agattgttgt ttccggcctt cagttaaac 11940
 tatcagtgtt tgacaggata tattggccgg taaacctaag agaaaaagagc gtttattaga 12000
 ataatcgat atttaaaagg gcgtgaaaag gtttattcctt cgtccatttg tatgtgcattg 12060
 ccaaccacag ggttcccca 12079

<210> 23

<211> 13002

<212> DNA

<213> Unknown

<220>

<223> pflanzlicher Expressionsvektor mit zwei Promotor-Terminator-Expressionskassetten

<400> 23

gatctggccgc cggccagcga gacgagcaag attggccgcc gcccggaaacg atccgacagc 60

gcgcccccagca caggtgcgcga ggcaaattgc accaacgcac acagcgccag cagaatgc 120

tagtggccgg tgacgtcggt cgagtgaacc agatcgcgca ggaggcccgg cagcacccggc 180
 ataatcaggc cgatgccgac agcgtcgagc ggcacagtgc tcagaattac gatcaggggt 240
 atgttgggtt tcacgtctgg cctccggacc agcctccgct ggtccgattt aacgcgcgga 300
 ttctttatca ctgataagtt ggtggacata ttatgtttat cagtgataaaa gtgtcaagca 360
 tgacaaagtt gcagccgaat acagtgtatcc gtgcggccct ggacctgttg aacgaggtcg 420
 gcgttagacgg tctgacgaca cgcaaactgg cgaaacgggtt gggggttcag cagccggcgc 480
 tttactggca cttcaggaac aagcgggcgc tgctcgacgc actggccgaa gccatgctgg 540
 cggagaatca tacgcattcg gtgccgagag ccgacgacga ctggcgctca tttctgatcg 600
 ggaatgccccg cagcttcagg caggcgctgc tcgcctaccg cgatggcgcg cgcatccatg 660
 ccggcacgacg accggggcgca ccgcagatgg aaacggccga cgcgcaagctt cgcttcctct 720
 gcgaggccggg tttttcggcc ggggacgccc tcaatgcgt gatgacaatc agctacttca 780
 ctgttggggc cgtgcttgag gaggcaggccg ggcacagcga tgccggcgag cgccggccgca 840
 ccgttgaaca ggctccgctc tcgcccgtgt tgccggccgc gatagacgcc ttgcacgaag 900
 ccgggtccgga cgcagcggttc gagcaggac tcgcgggtat tgtcgatgga ttggcgaaaa 960
 ggaggctcgt tgcaggaac gttgaaggac cgagaaagggt tgacgattta tcaggaccgc 1020
 tgccggaaacgca caacccactc actacagcag agccatgttag acaacatccc ctcccccttt 1080
 ccacccgcgtc agacgccccgt agcagccgc tacgggcttt ttcatgcctt gcctagcgt 1140
 ccaaggctca cggccgcgcgt cggcctctctt ggcggccttc tggcgctctt ccgttcctc 1200
 gctcaactgac tcgctgcgtc cggtcgttcg gctgcggcga gcccgtatcag ctcaactcaaa 1260
 ggcggtaata cggttatcca cagaatcagg ggataacgcga ggaaagaaca tgtgagcaaa 1320
 aggccagcaa aaggccagga accgtaaaaaa ggccgcgttg ctggcggttt tccataggct 1380
 ccgcggccctt gacgagcattc acaaaaatcg acgctcaagt cagaggtggc gaaacccgac 1440
 aggactataa agataccagg cgtttcccccc tggaaagctcc ctcgtgcgtc ctccctgttcc 1500
 gaccctgccc cttaccggat acctgtccgc ctttctccct tcgggaagcgt tggcggtttt 1560
 cccgtgcata accctgcttc ggggtcatta tagcgatttt ttccggatataat ccattctttt 1620
 tcgcacgata tacaggattt tgccaaagggt ttcgtgtaga ctttcttgg tgcgttcaac 1680
 ggcgtcagcc gggcaggata ggtgaagtag gcccacccgc gagcgggtgt tccttcttca 1740
 ctgtccctta ttgcacccgtc gccgtgcata acggaaatcc tgctctgcga ggctggccgg 1800
 ctaccgcccgg cgtaacagat gagggcaagc ggatggctga tgaaaccaag ccaaccagga 1860

agggcagccc acctatcaag gtgtactgcc ttccagacga acgaagagcg attgaggaaa 1920
 aggcggcggc ggccggcatg agcctgtcgg cctacctgct ggccgtcggc cagggctaca 1980
 aaatcacggg cgtcgtggac tatgagcagc tcccgagct ggccgcata aatggcgacc 2040
 tgggcccctt gggggcctg ctgaaactct ggctcaccga cgaccgcgc acggcgccgt 2100
 tcggtgatgc cacgatcctc gcccgtcgg cgaagatcga agagaagcag gacgagctt 2160
 gcaaggcat gatgggcgtg gtccgcgcga gggcagagcc atgactttt tagccgctaa 2220
 aacggccggg ggggtgcgcgt gattgccaag cacgtccccca tgccgtccat caagaagagc 2280
 gacttcgcgg agctggtgaa gtacatcacc gacgagcaag gcaagaccga gcgccttgc 2340
 gacgctcacc gggctggttg ccctcgccgc tgggctggcg gccgtctatg gccctgcaaa 2400
 cgccgcagaa acgcccgtcga agccgtgtgc gagacaccgc ggccgcggc gttgtggata 2460
 cctcgcggaa aacttggccc tcactgacag atgaggggcg gacgttgaca cttgagggc 2520
 cgactcaccg ggcgcggcgt tgacagatga gggcaggct cgatttcggc cggcgacgtg 2580
 gagctggcca gcctcgcaaa tcggcggaaaa cgcctgattt tacgcgagtt tcccacagat 2640
 gatgtggaca agcctggggtaa agtgcctt gccgtattga cacttgaggg gcgcgactac 2700
 tgacagatga ggggcgcgat cttgacact tgagggcag agtgcgtgaca gatgagggc 2760
 gcacctattt acatttgagg ggctgtccac aggcaaaaa tccagcattt gcaagggttt 2820
 ccgcggcgtt ttcggccacc gctaacctgt cttttaccc gttttaaac caatatttat 2880
 aaaccttgtt tttaaccagg gtcgcctt gtgcgcgtga ccgcgcacgc cgaaggggg 2940
 tgccccccct ttcgcgaccc tcccgcccg ctaacgcggg cttccatcc ccccaaggggc 3000
 tgcccccctc ggccgcgaac ggccctcacc caaaaatggc agcgcgtggca gtccttgc 3060
 ttgcgggat cggggcagta acgggatggg cgatcagccc gagcgcgacg cccggaaagca 3120
 ttgacgtgcc gcaggtgctg gcatcgacat tcagcgcacca ggtgcgggc agtgcggc 3180
 gcggcctggg tggcggcctg cccttcaattt cggccgtcgg ggcattcagc gacttcattgg 3240
 cggggccggc aatttttacc ttggcattt ttggcatagt ggtgcgggt gccgtgtcg 3300
 tgttcggggg tgcgataaac ccagcgaacc attttaggtg ataggtaaga ttataccgag 3360
 gtatgaaaac gagaattgga ctttacaga attactctat gaagcgcacat attaaaaag 3420
 ctaccaagac gaagaggatg aagaggatga ggaggcagat tgccttgcatttgc 3480
 tactgataag ataatatatc ttatatacg aagatatcgc cgtatgtaa gatttcagg 3540
 ggcaaggcat aggccgcgcg cttatcaata tatctataga atggcggaaaact 3600

tgcattggact aatgcttgaa acccaggaca ataaccttat agcttgtaaa ttctatcata 3660
attggtaat gactccaact tattgatagt gtttatgtt cagataatgc ccgatgactt 3720
tgtcatgcag ctccaccgat tttgagaacg acagcgactt ccgtcccagc cgtgccaggt 3780
gctgcctcag attcaggtta tgccgctcaa ttgcgtgcgt atatcgcttg ctgattacgt 3840
gcagcttcc ctcaaggcgg gattcataaca gcggccagcc atccgtcatc catatcacca 3900
cgtcaaaggg tgacagcagg ctcataagac gccccagcgt cgccatagtg cgttcaccga 3960
atacgtgcgc aacaaccgtc ttccggagac tgtcatacgc gtaaaacagc cagcgctggc 4020
gcgatttagc cccgacatag ccccactgtt cgtccatttc cgccagacg atgacgtcac 4080
tgccggctg tatgcgcgag gttaccgact gcggccttag ttttttaagt gacgtaaaat 4140
cgtgttgagg ccaacgccc taatgcgggc ttttgcggg catccaaacgc cattcatggc 4200
catatcaatg attttctggc gcttaccggg ttgagaagcg gtgttaagtga actgcagttg 4260
ccatgtttta cggcagttag agcagagata gcgcgtatgt ccggcgggtgc ttttgcgtt 4320
acgcaccacc cctgtcgttag ctgaacagga gggacagctg atagacacag aagccactgg 4380
agcacctcaa aaacaccatc atacactaaa tcagtaagtt ggcagcatca cccataattg 4440
tggttcaaa atcggctccg tcgatactat gttatacgcc aactttgaaa acaactttga 4500
aaaagctgtt ttctggatt taagttta gaatgcagg aacagtgaat tggagttcgt 4560
cttggtaaa ttagcttctt ggggtatctt taaatactgt agaaaagagg aaggaaataa 4620
taaatggcta aaatgagaat atcaccggaa ttgaaaaaac tgatcgaaaa ataccgctgc 4680
gtaaaagata cggaaaggaat gtctcctgct aaggtatata agctgggtgg agaaaatgaa 4740
aacctatatt taaaatgac ggacagccgg tataaaggga ccacctatga tgtggAACGG 4800
gaaaaggaca tgatgtatg gctggaaagga aagctgcctg ttccaaaggt cctgcacttt 4860
gaacggcatg atggctggag caatctgctc atgagtggagg ccgatggcgt ctttgcgt 4920
gaagagtatg aagatgaaca aagccctgaa aagattatcg agctgtatgc ggagtgcac 4980
aggcttttc actccatcga catatcgat tgtccctata cgaatagctt agacagccgc 5040
ttagccaat tggattactt actgaataac gatctggccg atgtggattg cgaaaactgg 5100
gaagaagaca ctccattna agatccgcgc gagctgtatg atttttaaa gacggaaaag 5160
cccgaaagagg aacttgtctt ttcccacggc gacctgggag acagcaacat ctttgcgt 5220
gatggcaaag taagtggctt tattgatctt gggagaagcg gcagggcggc caagtggat 5280
gacattgcct tctgcgtccg gtcgatcagg gaggatatcg gggaaagaaca gtatgcgag 5340

ctatTTTTG acttactggg gatcaaggct gattgggaga aaataaaata ttatattta 5400
 ctggatgaat tggTTTtagta cctagatgtg gcgcAACGAT gcccggcaca agcaggAGCG 5460
 caccgacttc ttccgcATCA agtgtttgg ctctcaggCC gaggcccACG gcaagtATTT 5520
 gggcaaggGGG tcgctggtat tcgtgcAGGG caagattCgg aataccaAGT acgagaAGGA 5580
 cggccAGACG gtctacggGA ccgacttcat tgCCGATAAG gtggattATC tggacacCAA 5640
 ggcaccAGGC gggTCAAATC aggaataAGG gcacATTGCC ccggcGTGAG tcggggCAAT 5700
 cccgcaagGA gggTGAATGA atcggacGTT tgaccggAAAG gcatacAGGC aagaACTGAT 5760
 cgacgcggGG tttccgCCG aggatGCCGA aaccatCGCA agccgcACCG tcATGCGTGC 5820
 gccccGCGAA acTTCCAGT ccgtcgGCTC gatggTCCAG caagCTACGG ccaAGATCGA 5880
 gCGCGACAGC gtGCAACTGG ctccccCTGC CCTGCCGCG CcatCGGCCG ccgtggAGCG 5940
 ttCGCGTCGT ctCGAACAGG aggCGGcAGG ttggcGAAG tcgatGACCA tcgacacCGC 6000
 aggaactatG acgaccaAGA agcgaaaaAC CGCCGGCGAG gacCTGGCAA aacaggTCAG 6060
 cgaggCCAAG caggCCGCGT tgctGAAACA cacGAAGCAG cAGATCAAGG aaATGCAAGT 6120
 ttccTTGTTc gatattGCGC cgtggCCGGA cacGATGCGA gCGATGCCAA acgacacGGC 6180
 ccgCTCTGCC ctgttCACCA CGCGCAACAA gaaaATCCCG CGCGAGGCGC tgcaAAACAA 6240
 ggtcattttc cacgtcaACA aggacgtGAA gatcacCTAC accggcGTcG agctgcggGC 6300
 cgacgatGAC gaactGGTGT ggcAGCAGGT gttggAGTAC gCGAACGCA CCCCTATCGG 6360
 cgagCCGATC acTTCAcGT tctacGAGCT ttGCCAGGAC ctggGCTGGT cgatcaATGG 6420
 ccggTATTAC acgaAGGCCG aggaATGCT gtcgCGCCTA caggCGACGG CGATGGGCTT 6480
 cacgtccGAC CGCttGGGC acTtGGAATC ggtgtcgCTG ctgcACCGCT tccgcgtcCT 6540
 ggaccgtggc aagaaaACGT cccgttgcca ggtcctGATC gacgaggAAA tcgtcgTGT 6600
 gtttgcTggc gaccactACA cgaaATTcat atgggagaAG taccGCAAGC tgTCGCCAC 6660
 ggcccGACGG atgttcGACT attcAGCTC GcaccGGGAG ccgtacCCGC tcaAGCTGGA 6720
 aacTTCCGC ctCATGTGCG gatcgGATTc cacCCGCGTG aagaAGTGGC GCGAGCAGGT 6780
 cggcGAAGCC tgcGAAGAGT tgcGAGGcAG cggcctGGTG gaacacGCT gggtaATGA 6840
 tgacCTGGTG cattGCAAAC gctAGGGCCT tgtggggTCA gttccGGCTG ggggttcAGC 6900
 agccAGCGCT ttactGGCAT ttcaGGAACA agcGGGcACT gctcgACGCA cttgcTTCGC 6960
 tcAGTATCGC tcgggACGCA cggcGCGCTC tacGAACTGC cgataAAACAG aggATTAaaa 7020
 ttgacaATTG tgattaAGGC tcagattcGA cggcTTGGAG cggccGACGT gcaggattTC 7080

cgcgagatcc gattgtcgcc cctgaagaaa gctccagaga tggtcgggtc cgtttacgag 7140
cacgaggaga aaaagccccat ggaggcggtc gctgaaacgggt tgcgagatgc cgtggcattc 7200
ggcgccctaca tcgacggcga gatcattggg ctgtcgggtct tcaaacagga ggacggccccc 7260
aaggacgctc acaaggcgca tctgtccggc gttttctgtgg agcccgaaca gcgaggccga 7320
ggggtcgccc gtatgctgct gccccgggtt ccggccgggtt tattgctctgatgatcg 7380
cgacagattc caacgggaat ctggtgatg cgcatacttca tcctcggcgc acttaatatt 7440
tcgctattct ggagcttgtt gtttatttctg gtctaccgcc tgccggccgg ggtcgcccgg 7500
acggtaggcg ctgtgcagcc gctgatggtc gtgttcatct ctgcgcgtct gctaggtac 7560
ccgatacgt tcatggcggt cctggggct atttgcggaa ctgcggcggtt ggccgtgtt 7620
gtgttgcacac caaacgcagc gctagatctt gtcggcgac cagcgggcct ggccggggcg 7680
tttccatgg eggtcgaaac cgtgctgacc cgcactggc aaccccccgt gcctctgctc 7740
acctttaccg cctggcaact ggccggccggaa ggacttctgc tcgttccagt agcttttagt 7800
tttgatccgc caatcccgat gcctacagga accaatgttc tcggcctggc gtggctcgcc 7860
ctgatcgag cgggtttaac ctacttctt tggttccggg ggatctcgcg actcgaacct 7920
acagttgttt ctttactggg ctttctcagc cccagatctg gggtcgatca gccggggatg 7980
catcaggccg acagtcggaa ctccgggtcc cgcacctgtt ccattcgggtg agcaatggat 8040
aggggagttt atatcgtaa cgttcacttc taaaagaaaata gcgcactca gcttcctcag 8100
cggtttatc cagcgatttc ctattatgtc ggcatagttc tcaagatcga cagcctgtca 8160
cggttaagcg agaaatgaat aagaaggctg ataattcggaa tctctgcgag ggagatgata 8220
tttgatcaca ggcagcaacg ctctgtcatc gttacaatca acatgctacc ctccgcgaga 8280
tcatccgtgt ttcaaaacccg gcagcttagt tgccgttctt ccgaatagca tcggtaacat 8340
gagcaaagtc tgccgcctta caacggctct cccgctgacg ccgtcccgga ctgatgggct 8400
gcctgtatcg agtggtgatt ttgtgcccag ctgcgggtcg gggagctgtt ggctggctgg 8460
tggcaggata tattgtggtg taaacaaatt gacgcttaga caacttaata acacattgcg 8520
gacgtttta atgtactggg gtggtttttc ttttaccag tgagacgggc aacagctgat 8580
tgcccttac cgcctggccc tgagagatgt gcagcaagcg gtccacgctg gtttgcggcca 8640
gcagggaaaa atcctgtttt atggtggttc cggaaatcgcc aaaatccctt ataaatcaaa 8700
agaatagccc gagataggggt tgagtgttgc tccagttgg aacaagagtc cactattaaa 8760
gaacgtggac tccaaacgtca aaggcgaaaa aaccgtctat cagggcgatg gcccactacg 8820

tgaaccatca cccaaatcaa gtttttggg gtcgaggtgc cgtaaagcac taaatcgaa 8880
 ccctaaaggg agccccgat ttagagctt acggggaaag ccggcgaacg tggcgagaaa 8940
 ggaagggaag aaagcgaaag gagcggcgc cattcaggct gcgcactgt tgggaaggc 9000
 gatcggtgcg ggccttctcg ctattacgcc agctggcgaa agggggatgt gctgcaaggc 9060
 gattaagttg ggtaacgcca gggtttccc agtcacgacg ttgtaaaacg acggccagt 9120
 aattaattcc catcttgaaa gaaataatagt ttaaatattt attgataaaa taacaagtca 9180
 ggtattatag tccaagcaaa aacataaatt tattgatgca agtttaaatt cagaatatt 9240
 tcaataactg attatatcag ctggtacatt gccgttagatg aaagactgag tgcgatatta 9300
 tgtgtataac ataaattgat gatatacgta gcttagctca tcgggggatc cgtcgaagct 9360
 agcttggtc ccgctcgaaa gaactcgtca agaaggcgat agaaggcgat gcgcgtcgaa 9420
 tcgggagcgg cgataccgta aagcacgagg aagcggtcag cccattcgcc gccaagctct 9480
 tcagcaatat cacggtagc caacgctatg tcctgatagc ggtccgccc acccagccgg 9540
 ccacagtcga tgaatccaga aaagcgccca tttccacca tgatattcgg caagcaggca 9600
 tcgcatggg tcacgacgag atcctcgccg tcggcatgc gcgccttgcg cctggcgaac 9660
 agttcggtcg ggcgcggccc ctgatgtct tcgtccagat catcctgatc gacaagaccg 9720
 gttccatcc gagtacgtgc tcgctcgatg cgatgttcg cttgggtggc gaatggcgag 9780
 gtagccggat caagcgtatg cagccgcgc attgcatcag ccatgatgga tactttctcg 9840
 gcaggagcaa ggtgagatga caggagatcc tgcccccggca cttcgcccaa tagcagccag 9900
 tccctcccg cttcagtgac aacgtcgagc acagctcgdc aaggaacgccc cgtcgtggcc 9960
 agccacgata gcccgcgtgc ctcgtcctgc agttcattca gggcacccgga caggtcggtc 10020
 ttgacaaaaaa gaaccggcgccc cccctgcgtc gacagccgga acacggcgcc atcagagcag 10080
 ccgattgtct gttgtgccc gtcatagccg aatagcctct ccacccaagc ggccggagaa 10140
 cctgcgtgca atccatcttgc ttcaatccaa gctccatgg gcccctgact agagtcgaga 10200
 tctggattga gagtgaatat gagactctaa ttggataccg agggaaattt atggaacgtc 10260
 agtggagcat tttgacaag aaatatttgc tagctgatag tgaccttagg cgactttga 10320
 acgcgcataa atggtttctg acgtatgtgc ttagctcatt aaactccaga aaccggcgcc 10380
 tgagttggctc cttcaacgtt gcggtctgt cagttccaaa cgtaaaacgg cttgtcccg 10440
 gtcatcgccg ggggtcataa cgtgactccc ttaattctcc gctcatgatc ttgatoccc 10500
 gcgccatcag atccttggcg gcaagaaagc catccagttt actttgcagg gcttcccaac 10560

cacttggagg atgcaataat gaagaaaact acaaattac atgcaactag ttatgcgtt 12360
 agtctatata atgaggattt tgcaatactt tcattcatac acactcacta agtttacac 12420
 gattataatt tcttcatacg cagcggatcc gatatcgggc cgctagcgt taaccctgt 12480
 ttaatgagat atgcgagacg cctatgatcg catgatattt gctttcaatt ctgttgtca 12540
 cgttgtaaaa aacctgagca tgttagctc agatcattac cgccggttc ggttcattct 12600
 aatgaatata tcacccgtta. ctatcgattt ttatgaata atattctccg ttcaatttac 12660
 tgattgtccg tcgacgaatt cgagctggc gcgcctctag aggatcgatg aattcagatc 12720
 ggctgagtgg ctccctcaac. gttgcggttc tgtagttcc aaacgtaaaa cggttgtcc 12780
 cgcgtcatcg gcgggggtca taacgtgact cccttaattc tccgctcatg atcagattgt 12840
 cgtttcccgc ctcaagttt aactatcagt gtttgacagg atatattggc gggtaaacct 12900
 aagagaaaaag agcgtttatt agaataatcg gatattaaa agggcgtgaa aaggtttac 12960
 cttcgtccat ttgtatgtgc atgccaacca cagggttccc ca 13002

<210> 24
 <211> 13905
 <212> DNA
 <213> Unknown

<220>
 <223> pflanzlicher Expressionsvektor mit drei
 Promotor-Terminator-Expressionskassetten

<400> 24
 gatctggcgc cggccagcga gacgagcaag attggccgcc gccccaaacg atccgacagc 60
 gcccggcaga cagggtgcgcga ggcaattgc accaacgcac acagcggcag cagaatgcca 120
 tagtggccgg tgacgtcggtt cgagtgaacc agatcgca ggaggccgg cagcacccgc 180
 ataatcaggc cgatgcccac agcgtcgagc ggcacagtgc tcagaattac gatcagggtt 240
 atgtgggtt tcacgtctgg cctccggacc agcctccgct ggtccgattt aacgcggcga 300
 ttcttatca ctgataagtt ggtggacata ttatgtttat cagtgataaa gtgtcaagca 360
 tgacaaagtt gcagcgaat acagtgtatcc gtgcggccct ggacctgttgc aacgagggtcg 420
 gcgtagacgg tctgacgaca cgcaaaactgg cgaaacgggtt gggggttcag cagccggcgc 480
 tttactggca ctcaaggaac aagcggccgc tgctcgacgc actggccgaa gccatgctgg 540
 cggagaatca tacgcattcg gtgcggagag ccgacgacga ctggcgctca tttctgtatcg 600
 ggaatgcggc cagcttcagg caggcgctgc tcgcctaccg cgatggcgccg cgcacatccatg 660
 cggcacgog accggccgcga ccgcagatgg aaacggccga cgcgacgtt cgcttcctct 720

gcgaggcggg ttttcggcc ggggacgccc tcaatgcgct gatgacaatc agctacttca 780
 ctgttgggc cgtgcttgag gagcaggccg gcgacagcga tgccggcgag cgccggggca 840
 ccgttgaaca ggctccgctc tcgcccgtgt tgccggccgc gatagacgcc .ttcgacgaag 900
 ccggtccgga cgcagcggttc gagcaggac tcgcggtgat tgtcgatgga ttggcgaaaa 960
 ggaggctcggt tgcaggaac gttgaaggac cgagaagg 1020
 tgccggagcg caacccactc actacagcag agccatgtag acaacatccc ctccccctt 1080
 ccaccgcgtc agacgcccgt agcagccgc tacggcttt ttcatgcctt gccctagcgt 1140
 ccaaggctca cggccgcgtc cggcctctt ggcggcccttc tggcgctctt ccgcttcctc 1200
 gctcaactgac togctgcgtc cggtcgttcg gctgcggcga gcggtatcag ctcactcaaa 1260
 ggccgttaata cggttatcca cagaatcagg ggataacgca ggaaaagaaca tgtgagcaaa 1320
 aggccagcaa aaggccagga accgtaaaaa gcccgcgttg ctggcggttt tccataggct 1380
 ccgccccctt gacgagcattc acaaaaatcg acgctcaagt cagaggtggc gaaacccgac 1440
 aggactataa agataaccagg cgtttcccccc tggaaagctcc ctcgtgcgtc ctccgttcc 1500
 gaccctgccc cttaccggat acctgtccgc ctttctccct tcgggaagcg tggcggttt 1560
 ccgctgcata accctgcttc ggggtcatta tagcgatttt ttcggatataat ccattttttt 1620
 tcgcacgata tacaggattt tgccaaagg 1680
 ggcgtcagcc gggcaggata ggtgaagtag gcccacccgc gagcgggtgt tcattttca 1740
 ctgtccctta ttgcacccgt gcggtgcgtca acgggaatcc tgctctgcga ggctggccgg 1800
 ctaccgcggg cgtaacagat gagggcaagc ggatggctga tgaaaccaag ccaaccagga 1860
 agggcagccc acctatcaag gtgtactgcc ttccagacga acgaagagcg attgaggaaa 1920
 aggcggccggc ggccggcatg agcctgtcgg cctacactgct ggccgtcggc cagggtaca 1980
 aaatcacggg cgtcgtggac tatgagcactc tccgcgagct ggccgcatac aatggcgacc 2040
 tggccgcctt gggcggcctg ctgaaactct ggctcaccga cgacccgcgc acggcgccgt 2100
 tcggtgatgc cacgatcctc gccctgtgg cgaagatcga agagaagcag gacgagctt 2160
 gcaaggtcat gatggcgttg gtccgcccga gggcagagcc atgactttt tagccgctaa 2220
 aacggccggg gggcgtggcgt gattgcaag cacgtccccca tgcgctccat caagaagagc 2280
 gacttcgcgg agctggtgaa gtacatcacc gacgagcaag gcaagaccga gccccttgc 2340
 gacgctcacc gggctggttg ccctcgccgc tgggtggcg gccgtctatg gccctgcataa 2400
 cgccgcagaa acgcccgtcga agccgtgtgc gagacaccgc gcccgcggc gttgtggata 2460

cctcgccgaa aacttggccc tcactgacag atgaggggcg gacgttgaca cttgagggc 2520
 cgactcaccc ggccgcggcgt tgacagatga ggggcaggct cgatttcggc cggcgacgtg 2580
 gagctggcca gcctcgcaaa tcggcgaaaa cgccctgattt tacgcgagtt tcccacagat 2640
 gatgtggaca agcctgggta taagtgcct gcggatttga cacttgaggg gcgcgactac 2700
 tgacagatga ggggcgcgat ctttgacact tgagggcag agtgcgtaca gatgagggc 2760
 gcaccttattt acatggagg ggctgtccac aggcaaaaa tccagcattt gcaagggtt 2820
 ccgcgcgtt ttcggccacc gctaaccctgt cttaaacctt gctttaaac caatattt 2880
 aaacccgtt tttaaccagg gctgcgcct gtgcgcgtga cgcgcacgc cgaaggggg 2940
 tgccccccct tatogaaccc tcccgccccg ctaacgcggg cctccatcc cccaggggc 3000
 tgcgccttc ggccgcgaac ggctcaccc caaaaatggc agcgctggca gtccttgcca 3060
 ttgcgggat cggggcagta acgggatggg cgatcagccc gagcgcgcacg cccggaaagca 3120
 ttgacgtgcc gcagggtgctg gcatcgacat tcagcgacca ggtgcggggc agtgagggcg 3180
 gcggcctggg tggcgccctg cccttcaactt cggccgtcgg ggcattcacg gacttcatgg 3240
 cggggccggc aatttttacc ttgggcattt ttggcatagt ggtgcgggtt gccgtgctcg 3300
 tgttcggggg tgcgataaac ccagcgaacc atttgaggtg ataggtaaga ttataccgag 3360
 gtatgaaaac gagaatttggc ctttacaga attactctat gaagcgccat attaaaaaag 3420
 ctaccaagac gaagaggatg aagaggatga ggaggcagat tgccttgaat atattgacaa 3480
 tactgataag ataatatatac ttttatatac aagatatcgc cgtatgtaa gatttcaggg 3540
 ggcaaggcat aggcagcgcg cttatcaata tatctataga atgggcaaag cataaaaact 3600
 tgcattggact aatgcttgc acccaggaca ataacctt atgcttgcataa ttctatcata 3660
 attggtaat gactccaact tattgatagt gttttatgtt cagataatgc ccgatgactt 3720
 tgtcatgcag ctccaccgat tttgagaacg acagcgactt cgcgtccacgc cgtgccaggt 3780
 gtcgcctcag attcaggta tggcgctcaa ttgcgtgcgt atatgccttgcgtt 3840
 gcagcttcc cttcaggcgg gattcataca gcccgcgcgc atccgtcatac catatcacca 3900
 cgtcaaaggc tgacagcagg ctcataagac gccccagcgt cgccatagtg cggtcaccga 3960
 atacgtgcgc aacaaccgtc ttccggagac tgtcatacgc gtaaaacagc cagcgctggc 4020
 gcgatttagc cccgacatag cccactgtt cgtccatttc cgcgcgcacg atgacgtcac 4080
 tgcccgctg tatgcgcgag gttaccgact gcccgcgtt gtttttaagt gacgtaaaat 4140
 cgtgttgagg ccaacgccc taatgcggc tggcgccgg catccaacgc cattcatggc 4200

cataatcaatg attttctggc gcgtaccggg ttgagaagcg gtgttaagtga actgcagttg 4260
ccatgtttta cggcagttag agcagagata gcgcgtatgt ccggcgggtgc ttttgccgtt 4320
acgcaccacc cccgtcagtag ctgaacagga gggacagctg atagacacag aagccactgg 4380
agcacctcaa aaacaccatc atacactaaa tcagtaagtt ggcagcatca cccataattg 4440
tggttcaaa atcggctccg tcgatactat gttatacgcc aactttgaaa acaacittga 4500
aaaagctgtt ttctggattt taaggtttta gaatgcaagg aacagtgaat tggagttcgt 4560
cttggtaaa tttagttctt ggggtatctt taaaactgt agaaaagagg aaggaaataa 4620
taaatggcta aaatgagaat atcaccggaa ttgaaaaaac tgatcgaaaa ataccgctgc 4680
gtaaaagata cggaaggaat gtctccgtct aaggtatata agctggtggg agaaaatgaa 4740
aacctatatt taaaatgac ggacagccgg tataaaggaa ccacctatga tgtggaacgg 4800
gaaaaggaca tgatgctatg gctggaagga aagctgcctg ttccaaaggt cctgcacttt 4860
gaacggcatg atggctggag caatctgctc atgagtgagg ccgatggcgt ccttgctcg 4920
gaagagtatg aagatgaaca aagccctgaa aagattatcg agctgtatgc ggagtgcac 4980
aggcttttc actccatcga catatcggt tgcctata cgaatagctt agacagccgc 5040
ttagccaat tggattactt actgaataac gatctggccg atgtggattt cgaaaactgg 5100
gaagaagaca ctccatattaa agatccgcgc gagctgtatg atttttaaa gacggaaaag 5160
cccgaagagg aacttgtctt ttcccacggc gacctgggag acagcaacat ctttgtaaa 5220
gatggcaaag taagtggctt tattgatctt gggagaagcg gcagggogga caagtggtat 5280
gacattgcct tctgcgtccg gtcgatcagg gaggatatcg gggaaaaaca gtatgtcgag 5340
ctatTTTt acttactgg gatcaagcct gattggaga aaataaaata ttatattta 5400
ctggatgaat tgTTTtagta cctagatgtg ggcacacgt gcccggcaca agcaggagcg 5460
caccgacttc ttccgcacca agtgtttgg ctctcaggcc gaggcccacg gcaagtatTT 5520
gggcaagggg tcgtcggtat tcgtgcaggg caagattcgg aataccaagt acgagaagga 5580
cgcccaagacg gtctacggga ccgacttcat tgccgataag gtggattatc tggacaccaa 5640
ggcaccaggc gggtaaaatc aggaataagg gcacattgcc ccggcgtcgt gtcggggcaat 5700
cccgcaagga gggtaatga atcggacgtt tgaccggaaag gcatacaggc aagaactgtat 5760
cgacgcgggg tttccgcgg aggatgccga aaccatcgca agccgcaccc tcgtcgatgc 5820
gccccgcgaa accttccagt ccgtcggtc gatggtccag caagctacgg ccaagatcga 5880
cgccgacacgc gtgcaactgg ctccccctgc cctgccccgcg ccacatcgccgc ccgtggagcg 5940

ttcgcgtcgt ctcgaacagg aggccggcagg tttggcgaag tcgatgacca tcgacacgcg 6000
 aggaactatg acgaccaaga agcgaaaaac cgccggcgag gacctggcaa aacaggtcag 6060
 cgaggccaag caggccgcgt tgctgaaaca cacgaagcag cagatcaagg aaatgcagct 6120
 ttccctgttc gatattgcgc cgtggccgga cacgatgcga gcgatgccaa acgacacggc 6180
 ccgctctgcc ctgttccacca cgcgcaacaa gaaaatcccg cgcgaggcgc tgcaaaacaa 6240
 ggtcattttc cacgtcaaca aggacgtgaa gatcacctac accggcgctg agctgcgggc 6300
 cgacgatgac gaactggtgt ggcagcaggt gttggagtac gcgaagcgca cccctatcg 6360
 cgagccgatc accttcacgt tctacgagct ttgccaggac ctgggcttgt cgatcaatgg 6420
 ccggtattac acgaaggccg aggaatgcct gtgcgccta cagggcgcgg cgatggcct 6480
 cacgtccgac cgcggtgggc acctggaatc ggtgtcgctg ctgcaccgct tccgcgtcct 6540
 ggaccgtggc aagaaaacgt cccgttgcca ggtcctgatc gacgaggaaa tcgtcgtgct 6600
 gtttgcgtgc gaccactaca cgaaattcat atgggagaag taccgcaagc tgtcgcgcac 6660
 ggccccacgg atgttcgact atttcagctc gcaccggag ccgtacccgc tcaagctgga 6720
 aaccttccgc ctcatgtgcg gatcgattc cacccgcgtg aagaagtggc gcgagcaggt 6780
 cggcgaagcc tgcgaaaggt tgcgaggcag cggcctggtg gaacacgcct gggtaatga 6840
 tgacctggtg cattgcaaacc gctaggccct tgggggtca gttccggctg ggggttcagc 6900
 agccagcgcct ttactggcat ttcaggaaca agcgggcact gtcgcacgc cttgcgtcgc 6960
 tcagtatcgc tcgggacgcga cggcgcgc tacaactgc cgataaaacag aggattaaaa 7020
 ttgacaattt tgattaaggc tcagattcga cggcttggag cggccgacgt gcaggatttc 7080
 cgcgagatcc gattgtcgcc cctgaagaaa gctccagaga tggtcggtc cgttacgag 7140
 cacgaggaga aaaagcccat ggaggcgatc gctgaacggc tgcgagatgc cgtggcattc 7200
 ggccgcctaca tcgacggcga gatcatggg ctgtcggtct tcaaacagga ggacggcccc 7260
 aaggacgcgc acaaggcgc tctgtccggc gttttcggtt agcccgaaca gcgaggccga 7320
 ggggtcgccg gtatgtcgct gctggcgatc cggcggtt tattgtcgat gatgtcgatc 7380
 cgacagattc caacgggaat ctgggtggatc cgcatttc tccatcgcc actaatatt 7440
 tcgctatttc ggagcttgcgtt gtttatttcg gtcattccgc tgccggcgatc ggtcgccgc 7500
 acggtaggcg ctgtgcagcc gctgatggtc gtgttcatct ctgcgcgtct gctaggtgc 7560
 ccgatacgat tgatggcggtt cctggggctt atttgcggaa ctgcggcgatc ggcgtgttg 7620
 gtgttgcacac caaacgcgc gctagatccct gtcggcgatc cagcggccct ggccggggcg 7680

gtttccatgg cgttcggAAC cgtgctgacc cgcaagtggc aaccccccgt gcctctgctc 7740
 acctttacgg cctggcaact ggccggccggaa ggacttctgc tcgttccagt agcttttagtg 7800
 tttgatccgc caatcccgat gcctacagga accaatgttc tcggcctggc gtggctcgcc 7860
 ctgatcgagg cgggtttaac ctacttcctt tggttccggg ggatctcgcg actcgaacct 7920
 acagttgttt ctttactggg ctttctcagc cccagatctg gggtcgatca gcccgggatg 7980
 catcaggccg acagtcggaa cttcgggtcc ccgacctgtc ccattcggtg agcaatggat 8040
 aggggagttg atatcgtaa cgttcacttc taaagaaaata gcgccactca gcttcctcag 8100
 cggcttatac cagcgatttc ctattatgtc ggcatacgatc tcaagatcga cagcctgtca 8160
 cggtaagcg agaaatgaat aagaaggctg ataattcggta tctctgcgag ggagatgata 8220
 tttgatcaca ggcagcaacg ctctgtcatc gttacaatca acatgttacc ctccgcgaga 8280
 tcatccgtgt ttcaaaccgg gcagcttagt tgccgttctt ccgaatagca tcggtaacat 8340
 gagcaaagtc tgccgcctta caacggctct cccgctgacg ccgtccggaa ctgtatggct 8400
 gcctgtatcg agtggtgatt ttgtgccgag ctgcccgtcg gggagctgtt ggctggctgg 8460
 tggcaggata tattgtggtg taaacaaatt gacgcttaga caacttaata acacattgcg 8520
 gacgaaaaatc atgtactggg gtggttttc ttttaccagg tgagacgggc aacagctgat 8580
 tgcccttcac cgcctggccc tgagagagtt gdagcaagcg gtccacgctg gttgccccca 8640
 gcaggcgaaa atccctgtttt atgggtggttc cggaaatcgcc aaaatccctt ataaatcaaaa 8700
 agaatagccc gagatagggt tgagtgttgt tccagttgg aacaagagtc cactattaaa 8760
 gaacgtggac tccaacgtca aaggcgaaa aaccgtctat cagggcgatg gcccaactacg 8820
 tgaaccatca cccaaatcaa gtttttggg gtcgaggtgc cgtaaagcac taaatcgaa 8880
 ccctaaaggag agcccccgtat ttagagcttg acggggaaag ccggcgaacg tggcgagaaa 8940
 ggaagggaaag aaagcgaaaag gagcggcgcc cattcaggct ggcgaactgt tgggaagggc 9000
 gatcggtgcg ggcctttcgc ctattacgccc agctggcgaa agggggatgt gctgcaaggc 9060
 gattaagttg gttaacgcca gggttttccc agtcacgacg ttgtaaaacg acggccagtg 9120
 aattaattcc catcttgaaa gaaatatagt ttaaatattt attgataaaa taacaagtca 9180
 ggtattatag tccaagcaaa aacataaatt tattgatgca agttaaatt cagaaatattt 9240
 tcaataactg attatatcag ctggtacatt gccgttagatg aaagactgag tgcgatatta 9300
 tgtgtataac ataaattgtat gatatacgta gcttagctca tcgggggatc cgtcgaagct 9360
 agcttggtc ccgctcagaa gaactcgtca agaaggcgat agaaggcgat ggcgtcgaa 9420

tcgggagccg cgataccgta aagcacgagg aagcggtcag cccattcgcc gccaagctct 9480
tcagcaatat cacggtagc caacgctatg tcctgatagc ggtccgccc acccagccgg 9540
ccacagtcga tgaatccaga aaagcggcca ttttccacca tgatattcgg caagcaggca 9600
tcgccatggg tcacgacgag atcctcgccg tcgggcatgc ggccttgag cctggcgaac 9660
agttcggtcg ggcggagccc ctgatgctct tcgtccagat catcctgatc gacaagaccg 9720
gcttccatcc gagtacgtgc tcgctcgatg cgatgttgc cttggtggtc gaatggcag 9780
gtagccggat caagcgtatg cagccgccc attgcatcag ccatgatgga tactttctcg 9840
gcaggagcaa ggtgagatga caggagatcc tgccccggca cttcgcccaa tagcagccag 9900
tcccttcccg cttcagtgac aacgtcgagc acagctgcgc aaggaacgccc cgctcggtcc 9960
agccacgata gccgcgtgc ctgcgtcgc agttcattca gggcaccggc caggtcggc 10020
ttgacaaaaaa gaaccggggcg cccctgcgt gacagccggc acacggccgc atcagagcag 10080
ccgattgtct gttgtgcggc gtcatalogg aatagcctct ccacccaagc ggccggagaa 10140
cctgcgtgca atccatcttg ttcaatccaa gctccatgg gccctcgact agagtcgaga 10200
tctggattga gagtgaatat gagactctaa ttggataccg aggggaattt atggAACGTC 10260
agtggagcat ttttgacaag aaatatttgc tagctgatag tgaccttagg cgactttga 10320
acgogcaata atggttctg acgtatgtgc ttagctcatt aaactccaga aaccggccgc 10380
tgagtggctc cttcaacggt gcgggtctgt cagttccaaa cgtaaaacgg cttgtccgc 10440
gtcatcgccg ggggtcataa cgtgactccc ttaattctcc gctcatgatc ttgatcccc 10500
gcgccatcag atccttggcg gcaagaaaagc catccagttt actttgcagg gttcccaac 10560
cttaccagag ggcgcggccag ctggcaattc cggttcgtt gctgtccata aaaccggcc 10620
gtctagctat cgccatgtaa gcccactgca agctacctgc tttctcttg cgcttgcgtt 10680
ttccctgtc cagatagccc agtagctgac attcatccgg ggtcagcacc gtttctgccc 10740
actggcttcc tacgtgttcc gtttcctta gcagcccttg cgccctgagt gtttgcggca 10800
gcgtgaagct tgcatgcctg caggtcgacg ggcgcggcggc ctccctcgagc aaatttacac 10860
attgccaacta aacgtctaaa ccctgtaat ttgttttgt tttactatgt gtgttatgta 10920
tttgatttgc gataaaatttt tatatttggt actaaattta taacacctt tatgctaacg 10980
tttgccaaaca ctttagcaatt tgcaaggtaa ttaattgatt ctaaaatttatttttcttct 11040
aaatacatat actaatcaac tggaaatgta aatatttgct aatatttctatctataggaga 11100
attaaagtga gtgaatatgg taccacaagg tttggagatt taattgttgc aatgctgcat 11160

ggatggcata tacaccaaac attcaataat tcttggaggat aataatggta ccacacaaga 11220
tttgagggtgc atgaacgtca cgtggacaaa agtttagta attttcaag acacaaatgt 11280
taccacacac aagtttgag gtgcattcat gnatgccctg tggaaagttt aaaaatattt 11340
tggaaatgat ttgcattggaa gccatgtgtaa accatgac atccacttgg aggatgcaat 11400
aatgaagaaa actacaaatt tacatgcaac tagttatgca tgtagtctat ataatgagga 11460
tttgcaata ctccattca tacacactca ctaagttta cacgattata atttcttcat 11520
agccagccca ccgcgggtggg cggccgcctg cagtctagaa ggcttcctgc tttaatgaga 11580
tatgcgagac gcctatgatc gcatgatatt tgcttcaat tctgttgtgc acgttgtaaa 11640
aaacctgagc atgtgttagct cagatccttac ccgcgggttt cggttcattc taatgaatat 11700
atcacccgtt actatcgat tttatgaat aatattctcc gttcaattta ctgattgtcc 11760
gtcgagcaaa ttacacatt gccactaaac gtctaaaccc ttgttaattt ttttgtttt 11820
actatgtgtt ttatgtattt gatttgcatt aaattttat atttggtaact aaatttataa 11880
cacctttat gctaacgttt gccaacactt agcaatttgc aagttgatta attgattcta 11940
aattttttt gtcttctaaa tacatataact aatcaactgg aaatgtaaat atttgctaatt 12000
atttctacta taggagaatt aaagtgagtg aatatggtaac cacaagggtt ggagatttaa 12060
ttgttcaat gctgcattggaa tggcatatac accaaacatt caataatttgc tgaggataat 12120
aatggtagcca cacaagattt gaggtgcattt aacgtcacgt ggacaaaagg ttttagtaatt 12180
tttcaagaca acaatgttac cacacacaag ttttgggttgc catgcattggaa tgccctgtgg 12240
aaagttaaa aatattttgg aaatgatttgc catggaaagcc atgtgtaaaa ccatgacatc 12300
cacttggagg atgcaataat gaagaaaact accaaatttac atgcaacttag ttatgcattgt 12360
agtctatata atgaggattt tgcaataactt tcattcatac acactcacta agttttacac 12420
gattataatt tcttcatacg cagcggatcc gatatcgccc ccgttagcgt taaccctgtct 12480
ttaatgagat atgcgagacg cctatgatcg catgatattt gctttcaattt ctgttgtgc 12540
cgttgtaaaa aacctgagca tggtagctc agatcattac ccgcgggtttc ggttcattct 12600
aatgaatata tcacccgttac ctatcgattt tttatgaata atattctccg ttcaatttac 12660
tgattgtccg tcgagcaat ttacacatttgc ccaactaaacg tctaaaccct tgtaatttgc 12720
ttttgttttta ctatgtgtgt tatgtatttgc atttgcata aattttataa tttggtaacta 12780
aatttataaactt acctttatgt ctaacgttttgc ccaacttac gcaatttgca agttgattaa 12840
ttgattctaa atttttttgc tcttcataat acatataacta atcaactggaa aatgtaaataa 12900

tttgctaata tttctactat aggagaatta aagtgagtga atatggtacc acaaggtttg 12960
 gagatttaat tggcataatg ctgcattggat ggcataataca ccaaacattc aataattctt 13020
 gaggataata atggtaccac acaagatttg aggtgcatac acgtcacgtg gacaaaaggt 13080
 ttagtaattt ttcaagacaa caatgttacc acacacaagt tttgagggtgc atgcattggat 13140
 gccctgtgga aagtttaaaa atattttgga aatgatttgc atggaagcca tgtgtaaaac 13200
 catgacatcc acttggagga tgcaataatg aagaaaacta caaatttaca tgcaactagt 13260
 tatgcatgta gtcttatataa tgaggattt gcaatacttt cattcataca cactcactaa 13320
 gttttacacd attataattt ctcatagcc agcagatctg ccggcatcga tcccgccca 13380
 tggcctgctt taatgagata tgcgagacgc ctatgatcgc atgatatttgc 13440
 ttttgtgcac gttgtaaaaa acctgaggcat gtgtagctca gatccttacc gcccgttgc 13500
 gttcattcta atgaatatat cacccgttac tatacgatattt ttatgaataa tatttcgt 13560
 tcaatttact gattgtccgt cgacgagctc ggccgcctc tagaggatcg atgaattcag 13620
 atcggctgag tggctcccttc aacgttgcgg ttctgtcagt tccaaacgta aaacggcttg 13680
 tccccgtca tcggccgggg tcataacgtg actcccttac ttctccgtc atgatcagat 13740
 tgtcgttcc cgcccttcagt taaaactatc agtgtttgac agatatttgc 13800
 cctaagagaa aagagcgttt attagaataa tcggatattt aaaagggcgt gaaaagggttt 13860
 atccttcgtc catttgcata tgcatgccaa ccacagggtt cccca 13905

<210> 25
 <211> 15430
 <212> DNA
 <213> Unknown

<220>
 <223> pflanz. Expressionsvektor mit zwei Promotor-Terminator-Expressionskassetten inseriert ist
Physcomitrella patens Elongase und Desaturase

<220>
 <221> CDS
 <222> (11543)..(12415)

<220>
 <221> CDS
 <222> (13313)..(14890)

<400> 25
 gatctggcgc cggccagcga gacgagcaag attggccgcc gcccgaaacg atccgacagc 60
 gcccggcagca caggtgcgcgca ggcaaattgc accaacgcac acagcgccag cagaatgcc 120

tagtggccgg tgacgtcggt cgagtgaacc agatcgcgca ggaggccccgg cagcacccggc 180
 ataattcaggc cgatgcccgc acgctcgagc gcgcacagtgc tcagaattac gatcagggtt 240
 atgttgggtt tcacgtctgg cctccggacc agcctccgct ggtccgattt aacgcgcggaa 300
 ttctttatca ctgataagtt ggtggacata ttatgtttat cagtataaaa gtgtcaagca 360
 tgacaaaagtt gcagccgaat acagtgtatcc gtgccgcctt ggacctgttg aacgaggctcg 420
 gcgttagacgg tctgacgaca cgcaaaactgg cgaaacgggtt gggggttcag cagccggcgc 480
 tttactggca cttcaggaac aagcgggcgc tgctcgacgc actggccgaa gccatgctgg 540
 cggagaatca tacgcattcg gtgccgagag ccgacgacga ctggcgtca ttctgtatcg 600
 ggaatgcccgg cagcttcagg caggcgctgc tcgcctaccg cgatggcgcg cgcacatccatg 660
 ccggcacgog accggggcgca ccgcagatgg aaacggccga cgcgcaagctt cgcttcctct 720
 gcgaggcggg ttttcggcc ggggacgccc tcaatgcgt gatgacaatc agctacttca 780
 ctgttggggc cgtgcttgag gagcaggccg ggcacagcga tgccggcgag cgccggggca 840
 ccgttgaaca ggctccgctc tcgcgcgtgt tgccggccgc gatagacgccc ttgcacgaag 900
 ccggtccggc cgcagcggtc gagcaggac tcgcgtgtat tgctgatgga ttggcgaaaa 960
 ggaggctcggt tgcaggaac gttgaaggac cgagaaagggt tgacgattga tcaggaccgc 1020
 tgccggagcg caacccactc actacacgcg agccatgttag acaacatccc ctcccccttt 1080
 ccaccgcgtc agacgcccgt agcagccgc tacggctttt ttcatgcctt gccctagcgt 1140
 ccaaggctca cggccgcgtc cggcctcttc ggcggccttc tggcgcttt ccgccttcctc 1200
 gctcaactgac tcgcgtcgct cggtcgttcg gctgcggcga gcggtatcag ctcactcaaaa 1260
 ggcggtaata cggttatcca cagaatcagg ggataacgcga ggaaagaaca tgcgatggaaa 1320
 aggccagcaa aaggccagga accgtaaaaa ggccgcgttg ctggcggttt tccataggct 1380
 ccgcggccctt gacgagcatc acaaaaatcg acgctcaagt cagaggtggc gaaacccgac 1440
 aggactataa agataccagg cgtttccccc tggaaagctcc ctgcgtcgct ctccgttcc 1500
 gaccctgccc cttaccggat acctgtccgc ctttctccct tcgggaagcg tggcgctttt 1560
 ccgctgcata accctgcttc ggggtcatta tagcgatttt ttccgtatata ccatcctttt 1620
 tcgcacgata tacaggattt tgccaaagggt ttcgtgtaga ctttccttgg tgcgtatcaac 1680
 ggcgtcagcc gggcaggata ggtgaagtag gcccacccgc gagcgggtgt tccttctca 1740
 ctgtccctta ttgcacactg gcgggtcgctca acgggaatcc tgctctgcga ggctggccgg 1800
 ctaccggccgg cgtaacagat gagggcaagc ggatggctga tgaaaccaag ccaaccagga 1860

agggcagccc acctatcaag gtgtactgcc ttccagacga acgaagagcg attgaggaaa 1920
 aggcggcggc gcccggcatg agcctgtcgg cctacctgct ggccgtcggc cagggctaca 1980
 aaatcacggg cgtcgtggac tatgagcacg tccgcgagct ggcccgcatc aatggcgacc 2040
 tgggccccct gggccgcctg ctgaaactct ggctcaccga cgaccgcgc acggcgccgt 2100
 tcggtgatgc cacgatcctc gccctgtgg cgaagatcga agagaagcag gacgagctt 2160
 gcaaggtcat gatgggcgtg gtccgcccga gggcagagcc atgactttt tagccgctaa 2220
 aacggccggg gggtgcgcggt gattgccaag cacgtccccca tgcgctccat caagaagac 2280
 gacttcgcgg agctggtgaa gtacatcacc gacgagcaag gcaagaccga ggcgccttgc 2340
 gacgctcacc gggctggttg ccctcgccgc tggctggcg gccgtctatg gccctgcaaa 2400
 cgcccgagaa acggcgctcga agccgtgtgc gagacaccgc ggccgcccggc gtttgtggata 2460
 cctcgcggaa aacttggccc tcactgacag atgaggggag gacgttgaca cttgagggc 2520
 cgactcaccg ggcgccccgt tgacagatga gggcaggct cgatttcggc cggcgacgtg 2580
 gagctggcca gcctcgcaaa tcggcgaaaa cgcctgattt tacgcgagtt tcccacagat 2640
 gatgtggaca agcctggggta taagtgcctt ggggtattga cacttgaggg ggcgcactac 2700
 tgacagatga gggcgcgat cttgacact tgagggcag agtgcgtaca gatgagggc 2760
 gcacctattt acatttgagg ggctgtccac aggcaaaaa tccagcattt gcaagggttt 2820
 ccgcggcttt ttcggccacc gctaacctgt ctttaacct gctttaaac caatatttt 2880
 aaaccttgtt tttaaccagg gctgcgcctt gtgcgcgtga cgcgcacgc cgaaggggg 2940
 tgccccccct ttcgcgaccc tcccgccccg ctaacgcggg ctcctccatcc cccaggggc 3000
 tgcccccctc ggccgcgaac ggctcaccac caaaaatggc agcgctggca gtccttgcca 3060
 ttgcgggat cggggcagta acgggatggg cgatcagccc gagcgcgacg cccgaaagca 3120
 ttgacgtgcc gcaggtgctg gcatcgacat tcagcgacca ggtgccgggc agtgagggc 3180
 cgccgcctggg tggcgccctg cccttacatt cggccgtcgg ggcattcacg gacttcatgg 3240
 cggggccggc aattttacc ttggcatttc ttggcatagt ggtcgccgggt gccgtgtcg 3300
 tgttcggggg tgcgataaac ccagcgaacc atttgaggtg ataggtaaga ttataccgag 3360
 gtatgaaaac gagaattgga ctttacaga attactctat gaagcgcctt attaaaaag 3420
 ctaccaagac gaagaggatg aagaggatga ggaggcagat tgccttgaat atattgacaa 3480
 tactgataag ataatatatc ttttatatacg aagatatcgc cgtatgtaa gatttcaggg 3540
 ggcaaggcat aggcagcgccg cttatcaata tatctataga atgggcaaag cataaaaact 3600

tgcatggact aatgcttcaa acccaggaca ataaccatat agcttgtaaa ttctatcata 3660
 attggtaat gactccaact tattgatgt gtttatgtt cagataatgc ccgatgactt 3720
 tgtcatgcag ctccacccat tttgagaacg acagcgactt ccgtcccagc cgtgccaggt 3780
 gctgcctcag attcaggta tgccgctcaa ttgcgtgcgt atatcgctt ctgattacgt 3840
 gcagcttcc cttcaggcgg gattcataca gcggccagcc atccgtcatc catatcacca 3900
 cgtcaaagg tgacagcagg ctcataagac gcccccagcgt cgccatagt cgttcaccga 3960
 atacgtgcgc aacaaccgtc ttccggagac tgtcatacgc gtaaaacagc cagcgctggc 4020
 gcgatttagc cccgacatag ccccactgtt cgtccatttc cgccgacacg atgacgtcac 4080
 tgccggctg tatgcgcgag gttacogact gcggcctgag ttttttaagt gacgtaaaat 4140
 cgtgtgagg ccaacgcccc taatgcgggc ttttgcggg catccaaacgc cattcatggc 4200
 catatcaatg attttctggt gcgtaccggg ttgagaagcg gtgttaagtga actgcagttg 4260
 ccatgtttta cggcagtgag agcagagata gcgtatgtt ccggcgggtc ttttgcgtt 4320
 acgcaccacc cctgtcgttag ctgaacagga gggacagctg atagacacag aagccactgg 4380
 agcacctcaa aaacaccatc atacactaaa tcagtaagtt ggcagcatca cccataattg 4440
 tggttcaaa atcggctccg tcgataactat gttatacgcc aactttgaaa acaactttga 4500
 aaaagctgtt ttctggatt taagttta gaatgcaagg aacagtgaat tggagttcgt 4560
 cttgttataa ttagcttctt ggggtatctt taaatactgt agaaaaagagg aaggaaataa 4620
 taaatggcta aaatgagaat atcaccggaa ttgaaaaaac tgatcgaaaa ataccgctgc 4680
 gtaaaaagata cggaaaggaat gtctcctgct aaggtatata agctggtggg agaaaaatgaa 4740
 aacatatatt taaaaatgac ggacagccgg tataaaggga ccacctatga tggaaacgg 4800
 gaaaaggaca tgatgtatg gctggaagga aagctgcctg ttccaaaggt cctgcacttt 4860
 gaacggcatg atggctggag caatctgctc atgagtgagg ccgtggcgt ctttgcgt 4920
 gaagagtatg aagatgaac aagccctgaa aagattatcg agctgtatgc ggagtgcac 4980
 aggcttttc actccatcga catatcgat tgccctata cgaatagctt agacagccgc 5040
 ttagccaat tggattactt actgaataac gatctggccg atgtggattg cgaaaaactgg 5100
 gaagaagaca ctccattaa agatccgcgc gagctgtatg atttttaaa gacggaaaag 5160
 cccgaagagg aacttgtctt ttcccacggc gacgtggag acagcaacat ctttgtaaa 5220
 gatggcaaag taagtggctt tattgatctt gggagaagcg gcagggcgg aagtggtat 5280
 gacattgcct tctgcgtccg gtcgatcagg gaggatatcg gggagaaca gtatgtcgag 5340

ctatTTTTG acttaactggg gatcaaggct gattgggaga aaataaaata ttatattta 5400
 ctggatgaat tgTTTtagta cctagatgtg gcgcAACGAT gcccggcaca agcaggAGCG 5460
 cacCGACTTC ttccgcATCA agtGTTTGG ctctcaggCC gagGCCAcG gcaAGTATTt 5520
 gggcaAGGGG tcgtcggtat tcgtcgAGGG caagattcgg aataccaAGT acgagaAGGA 5580
 cggccAGACG gtctacGGGA ccgacttcat tgccgataAG gtggattATC tggacacCAA 5640
 ggcaccAGGC gggTCAAATC aggaATAAGG gcacattGCC ccggcGTGAG tcggggCAAT 5700
 cccgcaAGGA gggTGAATGA atcggacGTT tgaccggAAAG gcatacAGGC aagaACTGAT 5760
 cgacgcGGGG tttccgCCG aggatGCCGA aaccatCGCA agccgcACCG tcAtgcgtgc 5820
 gccccGCGAA acTTCCAGT ccgtcgGCTC gatggtCCAG caagCTACGG ccaAGATCGA 5880
 ggcgcACAGC gtgcAACTGG ctccCTGC CCTGCCGCG ccAtCGGCCG ccgtggAGCG 5940
 ttcgctcgt ctCGAACAGG aggCGGcAGG ttggcGAAG tcgatgACCA tcgacacGCG 6000
 aggaACTATG acgACCAAGA agcgaaaaAC cggcggcGAG gacCTGGCAA aacaggTCAG 6060
 cgaggCCAAG caggCCGCGT tgctgAAACA cacGAAGCAG cAGATCAAGG AAATGcAGCT 6120
 ttcTTTgttc gatattgcgc cgtggccgGA cacgatgcga gcatgcCAA acgacacGGC 6180
 ccgctctGCC ctgttCACCA cgcgcAAACAA gaaaATCCCG cgcgaggcgc tgcaAAACAA 6240
 ggtcattttc cacgtcaACA aggacgtgAA gatcacCTAC accggcgtcg agctgcGGGc 6300
 cgacgatGAC gaactggTGT ggcAGCAGGT gttggAGTAC gcgaAGCGCA cccctatCGG 6360
 cgagccGATC acTTCAcGT tctacgAGt ttGCCAGGAC ctggGCTGgt cgatcaatGG 6420
 ccggTattAC acgaAGGCCG aggaATGCT gtcgcgcTA caggcGACGG cgatggcTT 6480
 cacgtccGAC cgcgttGGGc acctgGAATC ggtgtcgCTG ctgcACCGCT tccgcgtcCT 6540
 ggaccgtggc aagAAAACGT cccgttGCCA ggtcctGATC gacgaggAAA tcgtcgTGT 6600
 gtttgctggc gaccactACA cgAAATTCAt atgggAGAAG taccgcaAGC tgtcGCCGAC 6660
 ggccccGACGG atgttcGACT atttcAGCTC gcACCCGGAG cCGTACCCGC tcaAGCTGGA 6720
 aacCTTCCGC ctCATGTGCG gatcgGATTc cacCCGCGTG aagaAGTGGC gcgAGCAGGT 6780
 cggcgaAGCC tgcgaAGAGT tgcgAGGGcAG cggcctGGTG gaACACGCT gggtaATGA 6840
 tgacCTGGTG cattgcaAAAC gctaggGCCT tGtggggTCA gttccggCTG gggGTTcAGC 6900
 agccAGCGCT ttactGGCAT ttcAGGAACA agcGGGcACT gctcgACGCA ctGCTTCGC 6960
 tcAGTATCGC tcgggACGCA cggcgcGCTC tacGAACTGC cgataAAACAG aggATTAaaa 7020
 ttgacaATTG tgattaAGGC tcagattcGA cggcttGGAG cggccGACGT gcaggattc 7080

cgcgagatcc gattgtcggc cctgaagaaa gctccagaga tgttcgggta cgtttacgag 7140
cacgaggaga aaaagcccat ggaggcggtc gctgaacggt tgcgagatgc cgtggcattc 7200
ggcgctaca tcgacggcga gatcattggg ctgtcggtct tcaaacagga ggacggcccc 7260
aaggacgctc acaaggcgca tctgtccggc gtttcgtgg agcccgaaca gcgaggccga 7320
ggggtcgccg gtatgctgct gcgggcgtt ccggcgggtt tattgctcgatgatcg 7380
cgacagattc caacggaaat ctggtgatg cgcacatctca tcctcgccgc acttaatatt 7440
tcgctattct ggagcttggtt gtttatttcg gtctaccgccc tgccggccgg ggtcgccgc 7500
acggtaggcg ctgtgcagcc gctgatggcgtc gtgttcatct ctgcccgtct gctaggtac 7560
ccgatacgat tgcgtggcgtt cctggggctt atttgcggaa ctgcggcgtt ggcgtgttg 7620
gtgttgcacac caaacgcagc gctagatcct gtcggcgtcg cagcggccct ggccggggcg 7680
gtttccatgg cggtcgaaac cgtgctgacc cgcaagtggc aacctccgtt gcctctgctc 7740
acctttacccg cctggcaact ggccggccggaa ggacttctgc tcgttccagt agcttagtg 7800
tttgatccgc caatcccgat gcctacagga accaatgttc tcggccctggc gtggctcg 7860
ctgatcgag cgggtttaac ctacttcctt tggttccggg ggatctcgactc actcgAACCT 7920
acagttgttt ctttactggg ctttctcagc cccagatctg gggtcgatca gccggggatg 7980
catcaggccg acagtcggaa cttcgggtcc cccacctgtt ccattcggtg agcaatggat 8040
aggggagttt atatcgtaa cggtcacttc taaagaaaata ggcggccactca gcttcctcag 8100
cggtttatc cagcgatttc ctattatgtc ggcatacgatcc tcaagatcgatc cagcctgtca 8160
cggttaagcg agaaatgaat aagaaggctg ataattcgga. tctctgcgag ggagatgata 8220
tttgatcaca ggcagcaacg ctctgtcata gttacaatca acatgctacc ctccgcgaga 8280
tcatccgtgt ttcaaaacccg gcaagtttgt tgccgttctt ccgaatagca tcgtaacat 8340
gagcaaagtc tgccgcctta caacggctct cccgctgacg ccgtccggaa ctgatgggtt 8400
gcctgtatcg agtgggtatt ttgtggcggag ctgcccgtcg gggagctgtt ggctggctgg 8460
tggcaggata tattgtgggt taaacaaaatt gacgcttaga caacttaata acacattgcg 8520
gacgtttta atgtactggg gtgggttttc tttcaccag tgagacgggc aacagctgat 8580
tgcccttcac cgcctggccc tgagagagtt gcaagcaagcg gtccacgctg gttgccccca 8640
gcaggcgaaa atccgttttgc atgggtggttc cggaaatcgcc aaaatccctt ataaatcaaa 8700
agaatagccc gagatagggt tgagtgttgt tccagttgg aacaagatgc cactattaaa 8760
gaacgtggac tccaacgtca aaggcgaaa aaccgtctat cagggcgatg gcccactacg 8820

tgaaccatca cccaaatcaa gtttttggg gtcgaggtgc cgtaaagcac taaatcgaa 8880
 ccctaaaggg agcccccgat tttagagttt acggggaaag ccggcgaacg tggcgagaaa 8940
 ggaaggaaag aaagcgaaag gagcggcgac cattcaggct gcgcactgt tgggaaggc 9000
 gatcggtgcg ggcctttcg ctattacgcc agctggcgaa agggggatgt gctgcaaggc 9060
 gattaagttt ggtAACGCCA gggtttccc agtcacgacg ttgtaaaacg acggccagt 9120
 aattaattcc catcttgaaa gaaatatagt taaaatattt attgataaaa taacaagtca 9180
 ggtattatag tccaagcaaa aacataaatt tattgatgca agtttaaattt cagaatattt 9240
 tcaataactg attatatacg ctggtagatt gcggtagatg aaagactgag tgcatat 9300
 tgtgtataac ataaattttagt gatatacgta gcttagctca tcgggggatc cgctgaagct 9360
 agcttgggtc ccgctcagaa gaactcgta agaaggcgat agaaggcgat gcgcgtcgaa 9420
 tcgggagcgg cgataaccgta aagcacgagg aagcggtcag cccattcgcc gccaagctct 9480
 tcagcaatat cacgggtac caacgctatg tcctgatagc ggtccggcac acccagccgg 9540
 ccacagtcga tgaatccaga aaagcgccca ttttccadca tgatattcgg caagcaggca 9600
 tcgcatggg tcacgacgag atcctcgccg tcggcatgc gcgccttgcg cctggcgaac 9660
 agttcggtcg ggcgcggccc ctgtatgtct tcgtccagat catcctgatc gacaagaccg 9720
 gttccatcc gactacgtgc tcgctcgatg cgatgtttcg cttgggtggtc gaatggcag 9780
 gtagccggat caagcgatg cagccgcgc attgcatcag ccatgatgga tactttctcg 9840
 gcaggagcaa ggtgagatga caggagatcc tgccccggca cttcgcccaa tagcagccag 9900
 tccctcccg cttcagtgac aacgtcgagc acagctcgcc aaggaacgccc cgctgtggcc 9960
 agccacgata gcccgcgtgc ctgtatgtc agttcattca gggcacccggc caggtcggtc 10020
 ttgacaaaaaa gaaccggcgccc cccctgcgtc gacagccggc acacggcgcc atcagagcag 10080
 ccgattgtct gttgtggccca gtcatagccg aatagcctct ccacccaaagc ggccggagaa 10140
 cctgcgtgca atccatcttgc ttcaatccaa gctccatgg gcccctcgact agagtcgaga 10200
 tctggattga gactgatatac gagactctaa ttggataccg agggaaattt atgaaacgac 10260
 agtggagcat ttttgcacaaag aaatatttgc tagctgatag tgaccttagg cgactttga 10320
 acgcgcataa atggtttctg acgtatgtgc tttagctcatt aaactccaga aacccgcggc 10380
 tgagtggctc cttcaacgtt gcggtctgt cagttccaaa cgtaaaacgg cttgtcccgc 10440
 gtcatcgccg ggggtcataa cgtgactccc ttaattctcc gctcatgatc ttgatcccct 10500
 gcgccatcag atccttggcg gcaagaaagc catccagttt actttgcagg gttcccaac 10560

cttaccagag ggcgccccag ctggcaattc cggttcgctt gctgtccata aaaccgccc 10620
 gtcttagctat cgccatgtaa gcccaactgca agctacacctgc tttctctttg cgcttgcgtt 10680
 ttcccttgc cagatagccc agtagctgac attcatccgg ggtcagcacc gtttctgcgg 10740
 actggcttcc tacgtgttcc gcttcctta gcagcccttg cgccctgagt gcttgcggca 10800
 gcgtgaagct tgcatgcctg caggtcgacg ggcgcggcgg 10860
 attgccacta aacgtctaaa cccttgcata ttgttttgtt tttactatgt gtgttatgta 10920
 ttgtttgc gataaatttt tatatttggc actaaattta taacaccctt tatgctaacc 10980
 ttgcacaaca ctttagcaatt tgcaagttga ttaatttgatt ctaaatttatt tttgtcttct 11040
 aaatacatat actaatcaac tggaaatgta aatatttgct aatatttcta ctataggaga 11100
 attaaagtga gtgaatatgg taccacaagg ttggagatt taattgttgc aatgctgcat 11160
 ggatggcata tacaccaaacc attcaataat tcttgaggat aataatggta ccacacaaga 11220
 ttgtgggtgc atgaacgtca cgtggacaaa aggttttagta attttcaag acaacaatgt 11280
 taccacacac aagttttgag gtgcattgcat ggtgccctg tggaaagttt aaaaatattt 11340
 tggaaatgat ttgcattgaa gccatgtgta aaaccatgac atccacttgg aggatgcaat 11400
 aatgaagaaa actacaaatt tacatgcaac tagttatgca tgttagtctat ataatgagga 11460
 ttttgcata ctttcattca tacacactca ctaagttta cacgattata atttcttcat 11520
 agccagccca ccgcggtgaa aa atg gag gtc gtg gag aga ttc tac ggt gag 11572
 Met Glu Val Val Glu Arg Phe Tyr Gly Glu
 1 5 10
 ttg gat ggg aag gtc tcg cag ggc gtg aat gca ttg ctg ggt agt ttt 11620
 Leu Asp Gly Val Ser Gln Gly Val Asn Ala Leu Leu Gly Ser Phe
 15 20 25
 ggg gtg gag ttg acg gat acg ccc act acc aaa ggc ttg ccc ctc gtt 11668
 Gly Val Glu Leu Thr Asp Thr Pro Thr Lys Gly Leu Pro Leu Val
 30 35 40
 gac agt ccc aca ccc atc gtc ctc ggt gtt tct gta tac ttg act att 11716
 Asp Ser Pro Thr Pro Ile Val Leu Gly Val Ser Val Tyr Leu Thr Ile
 45 50 55
 gtc att gga ggg ctt ttg tgg ata aag gcc agg gat ctg aaa ccg cgc 11764
 Val Ile Gly Gly Leu Leu Trp Ile Lys Ala Arg Asp Leu Lys Pro Arg
 60 65 70
 gcc tcg gag cca ttt ttg ctc caa gct ttg gtg ctt gtg cac aac ctg 11812
 Ala Ser Glu Pro Phe Leu Leu Gln Ala Leu Val Leu Val His Asn Leu
 75 80 85 90
 ttc tgg ttt gcg ctc agt ctg tat atg tgc gtg ggc atc gct tat cag 11860

Phe Cys Phe Ala Leu Ser Leu Tyr Met Cys Val Gly Ile Ala Tyr Gln 95	100	105		
gct att acc tgg cgg tac tct ctc tgg ggc aat gca tac aat cct aaa Ala Ile Thr Trp Arg Tyr Ser Leu Trp Gly Asn Ala Tyr Asn Pro Lys 110	115	120	11908	
cat aaa gag atg gcg att ctg gta tac ttg ttc tac atg tct aag tac His Lys Glu Met Ala Ile Leu Val Tyr Leu Phe Tyr Met Ser Lys Tyr 125	130	135	11956	
gtg gaa ttc atg gat acc gtt atc atg ata ctg aag cgc agc acc agg Val Glu Phe Met Asp Thr Val Ile Met Ile Leu Lys Arg Ser Thr Arg 140	145	150	12004	
caa ata agc ttc ctc cac gtt tat cat cat tct tca att tcc ctc att Gln Ile Ser Phe Leu His Val Tyr His Ser Ser Ile Ser Leu Ile 155	160	165	170	12052
tgg tgg gct att gct cat cac gct cct ggc ggt gaa gca tat tgg tct Trp Trp Ala Ile Ala His His Ala Pro Gly Gly Glu Ala Tyr Trp Ser 175	180	185	12100	
gcg gct ctg aac tca gga gtg cat gtt ctc atg tat gcg tat tac ttc Ala Ala Leu Asn Ser Gly Val His Val Leu Met Tyr Ala Tyr Tyr Phe 190	195	200	12148	
ttg gct gcc tgc ctt cga agt agc cca aag tta aaa aat aag tac ctt Leu Ala Ala Cys Leu Arg Ser Ser Pro Lys Leu Lys Asn Lys Tyr Ile 205	210	215	12196	
ttt tgg ggc agg tac ttg aca caa ttc caa atg ttc cag ttt atg ctg Phe Trp Gly Arg Tyr Leu Thr Gln Phe Gln Met Phe Gln Phe Met Leu 220	225	230	12244	
aac tta gtg cag gct tac tac gac atg aaa acg aat gcg cca tat cca Asn Leu Val Gln Ala Tyr Tyr Asp Met Lys Thr Asn Ala Pro Tyr Pro 235	240	245	250	12292
caa tgg ctg atc aag att ttg ttc tac tac atg atc tcg ttg ctg ttt Gln Trp Leu Ile Lys Ile Leu Phe Tyr Tyr Met Ile Ser Leu Leu Phe 255	260	265		12340
ctt ttc ggc aat ttt tac gta caa aaa tac atc aaa ccc tct gac gga Leu Phe Gly Asn Phe Tyr Val Gln Lys Tyr Ile Lys Pro Ser Asp Gly 270	275	280		12388
aag caa aag gga gct aaa act gag tga tctagaaggc ctccctgcttt Lys Gln Lys Gly Ala Lys Thr Glu 285	290			12435
aatgagatat gcgagacgcc tatgatcgca tgatatttg tttcaattct gttgtgcacg				12495
ttgtaaaaaaaaa cctgagcatg tgtagctcag atccttaccg ccggttcgg ttcatctaa				12555
tgaatatatac acccgttact atcgtatccc tatgaataat attctccgtt caatctactg				12615
attgtccgtc gagcaaattt acacattgcc actaaacgtc taaaacccttg taatttgttt				12675

Asp Glu His Pro Gly Gly Ser Val Ile Ser Thr Tyr Phe Gly Arg Asp
 430 435 440

 ggc aca gat gtt ttc tct agt ttt cat gca gct tct aca tgg aaa att 13810
 Gly Thr Asp Val Phe Ser Ser Phe His Ala Ala Ser Thr Trp Lys Ile
 445 450 455

 ctt caa gac ttt tac att ggt gac gtg gag agg gtg gag ccc act cca 13858
 Leu Gln Asp Phe Tyr Ile Gly Asp Val Glu Arg Val Glu Pro Thr Pro
 460 465 470

 gag ctg ctg aaa gat ttc cga gaa atg aga gct ctt ttc ctg agg gag 13906
 Glu Leu Leu Lys Asp Phe Arg Glu Met Arg Ala Leu Phe Leu Arg Glu
 475 480 485

 caa ctt ttc aaa agt tcg aaa ttg tac tat gtt atg aag ctg ctc acg 13954
 Gln Leu Phe Lys Ser Ser Lys Leu Tyr Tyr Val Met Lys Leu Leu Thr
 490 495 500 505

 aat gtt gct att ttt gct gcg agc att gca ata ata tgt tgg agc aag 14002
 Asn Val Ala Ile Phe Ala Ala Ser Ile Ala Ile Ile Cys Trp Ser Lys
 510 515 520

 act att tca gcg gtt ttg gct tca gct tgt atg atg gct ctg tgt ttc 14050
 Thr Ile Ser Ala Val Leu Ala Ser Ala Cys Met Met Ala Leu Cys Phe
 525 530 535

 caa cag tgc gga tgg cta tcc cat gat ttt ctc cac aat cag gtg ttt 14098
 Gln Gln Cys Gly Trp Leu Ser His Asp Phe Leu His Asn Gln Val Phe
 540 545 550

 gag aca cgc tgg ctt aat gaa gtt gtc ggg tat gtg atc ggc aac gcc 14146
 Glu Thr Arg Trp Leu Asn Glu Val Val Gly Tyr Val Ile Gly Asn Ala
 555 560 565

 gtt ctg ggg ttt agt aca ggg tgg tgg aag gag aag cat aac ctt cat 14194
 Val Leu Gly Phe Ser Thr Gly Trp Trp Lys Glu Lys His Asn Leu His
 570 575 580 585

 cat gct gct cca aat gaa tgc gat cag act tac caa cca att gat gaa 14242
 His Ala Ala Pro Asn Glu Cys Asp Gln Thr Tyr Gln Pro Ile Asp Glu
 590 595 600

 gat att gat act ctc ccc ctc att gcc tgg agc aag gac ata ctg gcc 14290
 Asp Ile Asp Thr Leu Pro Leu Ile Ala Trp Ser Lys Asp Ile Leu Ala
 605 610 615

 aca gtt gag aat aag aca ttc ttg cga atc ctc caa tac cag cat ctg 14338
 Thr Val Glu Asn Lys Thr Phe Leu Arg Ile Leu Gln Tyr Gln His Leu
 620 625 630

 ttc ttc atg ggt ctg tta ttt ttc gcc cgt ggt agt tgg ctc ttt tgg 14386
 Phe Phe Met Gly Leu Leu Phe Phe Ala Arg Gly Ser Trp Leu Phe Trp
 635 640 645

 agc tgg aga tat acc tct aca gca gtg ctc tca cct gtc gac agg ttg 14434
 Ser Trp Arg Tyr Thr Ser Thr Ala Val Leu Ser Pro Val Asp Arg Leu
 650 655 660 665

ttg gag aag gga act gtt ctg ttt cac tac ttt tgg ttc gtc ggg aca Leu Glu Lys Gly Thr Val Leu Phe His Tyr Phe Trp Phe Val Gly Thr 670	675	680	14482	
gcg tgc tat ctt ctc cct ggt tgg aag cca tta gta tgg atg gcg gtg Ala Cys Tyr Leu Leu Pro Gly Trp Lys Pro Leu Val Trp Met Ala Val 685	690	695	14530	
act gag ctg atg tcc ggc atg ctg ctg ggc ttt gta ttt gta ctt agc Thr Glu Leu Met Ser Gly Met Leu Leu Gly Phe Val Phe Val Leu Ser 700	705	710	14578	
cac aat ggg atg gag gtt tat aat tcg tct aaa gaa ttc gtg agt gca His Asn Gly Met Glu Val Tyr Asn Ser Ser Lys Glu Phe Val Ser Ala 715	720	725	14626	
cag atc gta tcc aca cgg gat atc aaa gga aac ata ttc aac gac tgg Gln Ile Val Ser Thr Arg Asp Ile Lys Gly Asn Ile Phe Asn Asp Trp 730	735	740	745	14674
ttc act ggt ggc ctt aac agg caa ata gag cat cat ctt ttc cca aca Phe Thr Gly Gly Leu Asn Arg Gln Ile Glu His His Leu Phe Pro Thr 750	755	760	14722	
atg ccc agg cat aat tta aac aaa ata gca cct aga gtg gag gtg ttc Met Pro Arg His Asn Leu Asn Lys Ile Ala Pro Arg Val Glu Val Phe 765	770	775	14770	
tgt aag aaa cac ggt ctg gtg tac gaa gac gta tct att gct acc ggc Cys Lys His Gly Leu Val Tyr Glu Asp Val Ser Ile Ala Thr Gly 780	785	790	14818	
act tgc aag gtt ttg aaa gca ttg aag gaa gtc gcg gag gct gcg gca Thr Cys Lys Val Leu Lys Ala Leu Lys Glu Val Ala Glu Ala Ala Ala 795	800	805	14866	
gag cag cat gct acc acc agt taa gctagcgtta accctgcttt aatgagatat Glu Gln His Ala Thr Ser 810	815		14920	
gcgagacgcc tatgatcgca tgatattgc tttcaattct gttgtgcacg ttgtaaaaaa			14980	
cctgagcatg tgttagctcg atccttaccg ccggtttcgg ttcattctaa tgaatatatc			15040	
acccgttact atcgtatccc tatgaataat attctccgtt caatttactg attgtccgtc			15100	
gacgaattcg agctcggcgc gcctcttagag gatcgatgaa ttccagatcgg ctgagtggct			15160	
ccttcaacgt tgccgttctg tcaagttccaa acgtaaaacg gcttgcggc cgtcatccgc			15220	
gggggtcata acgtgactcc cttattctc cgctcatgtat cagattgtcg ttccccgcct			15280	
tcagttaaa ctatcgtgt ttgacaggat atattggcgg gtaaacctaa gagaaaagag			15340	
cgtttattag aataatcgga tattttaaag ggcgtaaaaa ggtttatcct tcgtccattt			15400	
gtatgtcat gccaaaccaca gggttccccca			15430	

<210> 26
<211> 290
<212> PRT
<213> Unknown

<400> 26
Met Glu Val Val Glu Arg Phe Tyr Gly Glu Leu Asp Gly Lys Val Ser
1 5 10 15
Gln Gly Val Asn Ala Leu Leu Gly Ser Phe Gly Val Glu Leu Thr Asp
20 25 30
Thr Pro Thr Thr Lys Gly Leu Pro Leu Val Asp Ser Pro Thr Pro Ile
35 40 45
Val Leu Gly Val Ser Val Tyr Leu Thr Ile Val Ile Gly Gly Leu Leu
50 55 60
Trp Ile Lys Ala Arg Asp Leu Lys Pro Arg Ala Ser Glu Pro Phe Leu
65 70 75 80
Leu Gln Ala Leu Val Leu Val His Asn Leu Phe Cys Phe Ala Leu Ser
85 90 95
Leu Tyr Met Cys Val Gly Ile Ala Tyr Gln Ala Ile Thr Trp Arg Tyr
100 105 110
Ser Leu Trp Gly Asn Ala Tyr Asn Pro Lys His Lys Glu Met Ala Ile
115 120 125
Leu Val Tyr Leu Phe Tyr Met Ser Lys Tyr Val Glu Phe Met Asp Thr
130 135 140
Val Ile Met Ile Leu Lys Arg Ser Thr Arg Gln Ile Ser Phe Leu His
145 150 155 160
Val Tyr His His Ser Ser Ile Ser Leu Ile Trp Trp Ala Ile Ala His
165 170 175
His Ala Pro Gly Gly Glu Ala Tyr Trp Ser Ala Ala Leu Asn Ser Gly
180 185 190
Val His Val Leu Met Tyr Ala Tyr Phe Leu Ala Ala Cys Leu Arg
195 200 205
Ser Ser Pro Lys Leu Lys Asn Lys Tyr Leu Phe Trp Gly Arg Tyr Leu
210 215 220
Thr Gln Phe Gln Met Phe Gln Phe Met Leu Asn Leu Val Gln Ala Tyr
225 230 235 240
Tyr Asp Met Lys Thr Asn Ala Pro Tyr Pro Gln Trp Leu Ile Lys Ile
245 250 255
Leu Phe Tyr Tyr Met Ile Ser Leu Leu Phe Leu Phe Gly Asn Phe Tyr
260 265 270

Val Gln Lys Tyr Ile Lys Pro Ser Asp Gly Lys Gln Lys Gly Ala Lys
275 280 285

Thr Glu
290

<210> 27
<211> 525
<212> PRT
<213> Unknown

<400> 27
Met Val Phe Ala Gly Gly Gly Leu Gln Gln Gly Ser Leu Glu Glu Asn
1 5 10 15

Ile Asp Val Glu His Ile Ala Ser Met Ser Leu Phe Ser Asp Phe Phe
20 25 30

Ser Tyr Val Ser Ser Thr Val Gly Ser Trp Ser Val His Ser Ile Gln
35 40 45

Pro Leu Lys Arg Leu Thr Ser Lys Lys Arg Val Ser Glu Ser Ala Ala
50 55 60

Val Gln Cys Ile Ser Ala Glu Val Gln Arg Asn Ser Ser Thr Gln Gly
65 70 75 80

Thr Ala Glu Ala Leu Ala Glu Ser Val Val Lys Pro Thr Arg Arg Arg
85 90 95

Ser Ser Gln Trp Lys Lys Ser Thr His Pro Leu Ser Glu Val Ala Val
100 105 110

His Asn Lys Pro Ser Asp Cys Trp Ile Val Val Lys Asn Lys Val Tyr
115 120 125

Asp Val Ser Asn Phe Ala Asp Glu His Pro Gly Gly Ser Val Ile Ser
100 115 140

Thr Tyr Phe Gly Arg Asp Gly Thr Asp Val Phe Ser Ser Phe His Ala
145 150 155 160

Ala Ser Thr Trp Lys Ile Leu Gln Asp Phe Tyr Ile Gly Asp Val Glu
165 170 175

Arg Val Glu Pro Thr Pro Glu Leu Leu Lys Asp Phe Arg Glu Met Arg
180 185 190

Ala Leu Phe Leu Arg Glu Gln Leu Phe Lys Ser Ser Lys Leu Tyr Tyr
125 200 205

Val Met Lys Leu Leu Thr Asn Val Ala Ile Phe Ala Ala Ser Ile Ala
210 215 220

Ile Ile Cys Trp Ser Lys Thr Ile Ser Ala Val Leu Ala Ser Ala Cys
225 230 235 240

Met Met Ala Leu Cys Phe Gln Gln Cys Gly Trp Leu Ser His Asp Phe
 245 250 255
 Leu His Asn Gln Val Phe Glu Thr Arg Trp Leu Asn Glu Val Val Gly
 260 265 270
 Tyr Val Ile Gly Asn Ala Val Leu Gly Phe Ser Thr Gly Trp Trp Lys
 275 280 285
 Glu Lys His Asn Leu His His Ala Ala Pro Asn Glu Cys Asp Gln Thr
 290 295 300
 Tyr Gln Pro Ile Asp Glu Asp Ile Asp Thr Leu Pro Leu Ile Ala Trp
 305 310 315 320
 Ser Lys Asp Ile Leu Ala Thr Val Glu Asn Lys Thr Phe Leu Arg Ile
 325 330 335
 Leu Gln Tyr Gln His Leu Phe Phe Met Gly Leu Leu Phe Phe Ala Arg
 340 345 350
 Gly Ser Trp Leu Phe Trp Ser Trp Arg Tyr Thr Ser Thr Ala Val Leu
 355 360 365
 Ser Pro Val Asp Arg Leu Leu Glu Lys Gly Thr Val Leu Phe His Tyr
 370 375 380
 Phe Trp Phe Val Gly Thr Ala Cys Tyr Leu Leu Pro Gly Trp Lys Pro
 385 390 395 400
 Leu Val Trp Met Ala Val Thr Glu Leu Met Ser Gly Met Leu Leu Gly
 405 410 415
 Phe Val Phe Val Leu Ser His Asn Gly Met Glu Val Tyr Asn Ser Ser
 420 425 430
 Lys Glu Phe Val Ser Ala Gln Ile Val Ser Thr Arg Asp Ile Lys Gly
 435 440 445
 Asn Ile Phe Asn Asp Trp Phe Thr Gly Leu Asn Arg Gln Ile Glu
 450 455 460
 His His Leu Phe Pro Thr Met Pro Arg His Asn Leu Asn Lys Ile Ala
 465 470 475 480
 Pro Arg Val Glu Val Phe Cys Lys Lys His Gly Leu Val Tyr Glu Asp
 485 490 495
 Val Ser Ile Ala Thr Gly Thr Cys Lys Val Leu Lys Ala Leu Lys Glu
 500 505 510
 Val Ala Glu Ala Ala Ala Glu Gln His Ala Thr Thr Ser
 515 520 525

<212> DNA
 <213> Unknown

<220>
 <223> pflanz. Expressionsvektor mit 3
 Promotor-Terminator- Expressionskassetten
 inseriert mit Physcomitrella Elongase + Desaturase
 + Phaeodactylum Desaturase

<220>
 <221> CDS
 <222> (11543) .. (12415)

<220>
 <221> CDS
 <222> (13313) .. (14890)

<220>
 <221> CDS
 <222> (15791) .. (17200)

<400> 28

gatctggcgc cggccagcga gacgagcaag attggccgcc gccccaaacg atccgacagc 60
 gcgcccagca caggtgcgca ggcaaattgc accaacgcac acagcgccag cagaatgcc 120
 tagtggccgg tgacgtcggt cgagtgaacc agatcgcgca ggaggcccg cagcacccgc 180
 ataatcaggc cgatgcccac agcgtcgagc ggcacagtgc tcagaattac gatcaggggt 240
 atgttgggtt tcacgtctgg cctccggacc agcctccgct ggtccgattg aacgcgcgga 300
 ttctttatca ctgataagtt ggtggacata ttatgtttat cagtgataaa gtgtcaagca 360
 tgacaaaagtt gcagccgaat acagtgtatcc gtgccgcct ggacctgttg aacgaggctg 420
 gcgttagacgg tctgacgaca cgcaaactgg cgaaacgggtt gggggttcag cagccggcgc 480
 ttactggca cttcaggaac aagcgggcgc tgctcgacgc actggccgaa gccatgctgg 540
 cggagaatca tacgcattcg gtgccgagag ccgacgacga ctggcgctca ttctgtatcg 600
 ggaatgcccg cagttcagg caggcgctgc tcgcctaccg cgatggcgcg cgcatccatg 660
 cccgcacgcg accggcgca ccgcagatgg aaacggccga cgccgcagctt cgcttcctct 720
 gcgaggcggg ttttcggcc gggacgccc tcaatgcgct gatgacaatc agctacttca 780
 ctgttgggc cgtgctttag gagcaggccg ggcacagcga tgccggcgag cgccggcga 840
 ccgttgaaca ggctccgctc tgcgcgtgt tgccggccgc gatagacgcc ttgcacgaag 900
 ccggtccgga cgcagcggttc gagcaggac tcgcggtgat tgtcgatgga ttggcggaaaa 960
 ggaggctcggt tgtcaggaac gttgaaggac cgagaaaggg tgacgattga tcaggaccgc 1020
 tgccggagcg caacccactc actacagcag agccatgtag acaacatccc ctcccccttt 1080

ccaccgcgtc agacgcccgt agcagccgc tacgggctt ttcatgcctt gccttagcgt 1140
 ccaaggctca cggcccgcgct cggcctctct ggccggccttc tggcgctttt ccgttcctc 1200
 gctactgac tcgctgcgt cggtcggtcg gctgcggcga gcggtatcag ctcactcaaa 1260
 ggccgtataa cggttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa 1320
 aggccagcaa aaggccagga accgtaaaaaa ggccgcgttg ctggcgcccc tccataggct 1380
 ccgccccctt gacgagcatc acaaaaatcg acgctcaagt cagaggtggc gaaaccgac 1440
 aggactataa agataaccagg cgtttcccccc tggaaagctcc ctcgtgcgt ctccctgttcc 1500
 gaccctgccc cttaccggat acctgtccgc ctttctccct tcgggaagcg tggcgctttt 1560
 ccgctgcata accctgcttc ggggtcatta tagcgatttt ttcggtatata ccattctttt 1620
 tcgcacgata tacaggattt tgccaaaggg ttcgtgtaga ctttccttgg tgtatccaac 1680
 ggcgtcagcc gggcaggata ggtgaagtag gcccacccgc gagcgggtgt tccttcttca 1740
 ctgtccctta ttgcacccctg gcggtgctca acgggaatcc tgctctgcga ggctggccgg 1800
 ctaccgcggc cgtaacagat gagggcaagc ggatggctga tgaaaccaag ccaaccagga 1860
 agggcagccc acctatcaag gtgtactgcc ttccagacga acgaagagcg attgaggaaa 1920
 aggccggccgc ggccggcatg agcctgtcgg cctacctgct ggccgtcggc cagggctaca 1980
 aaatcacggg cgtcgtggac tatgagcacg tccgcgagct ggcccgcatc aatggcgacc 2040
 tggccgcctt gggccgcctg ctgaaactct ggctcaccga cgaccgcgc acggcgccgt 2100
 tcggtgatgc cacgatcctc gcccgttgg cgaagatcga agagaagcag gacgagctt 2160
 gcaaggatcat gatgggcgtg gtccgcggcga gggcagagcc atgacttttt tagccgctaa 2220
 aacggccggg gggtgccgt gattgccaag cacgtccccca tgcgcctccat caagaagagc 2280
 gacttcgcgg agctggtaa gtacatcacc gacgagcaag gcaagaccga gcgcctttgc 2340
 gacgctcacc gggctggttt ccctcgccgc tggcgtggcg gacgtctatg gcccgcctt 2400
 cgccgcagaa acgcccgtcga agccgtgtgc gagacaccgc ggccgcggc gttgtggata 2460
 cctcgccggaa aacttggccc tcactgacag atgagggcg gacgttgaca cttgaggggc 2520
 cgactcaccc ggccggcggt tgacagatga gggcaggct cgatttcggc cggcgacgtg 2580
 gagctggcca gcctcgcaaa tcggcgaaaa cgccctgattt tacgcgagtt tcccacagat 2640
 gatgtggaca agcctggggta taagtgcctt gcggtattga cacttgagggg gcgactac 2700
 tgacagatga gggcgcgat ctttgacact tgagggcgag agtgctgaca gatgaggggc 2760
 gcacctattt acatttgagg ggctgtccac aggcagaaaa tccagcattt gcaagggttt 2820

ccgcccgttt ttcggccacc gctaacctgt ctttaacct gctttaaac caatatttat 2880
aacaccttgtt tttaaccagg gctgcgcctt gtgcgcgtga ccgcgcacgc cgaagggggg 2940
tgccccccct tctcgaaccc tccccggcccg ctaacgcggg cctcccatcc ccccaaggggc 3000
tgcgccccctc ggccgcgaac ggccctcaccc caaaaatggc agcgctggca gtccttgcca 3060
ttgcccggat cggggcagta acgggatggg cgatcagccc gagcgcgcacg cccggaagca 3120
ttgacgtgcc gcagggtctg .gcatcgacat tcagcgacca ggtgccggc agtgagggcg 3180
gcggcctggg tggoggcctg cccttactt cggccgtcgg ggcattcacg gacttcatgg 3240
cggggccggc aatttttacc ttgggcattc ttggcatagt ggtcgcgggt gccgtgctcg 3300
tgttcggggg tgcgataaaac ccagcgaacc atttgagggtg ataggtaaga ttataccgag 3360
gtatgaaaac gagaattgga .cctttacaga attactctat gaagcgccat attaaaaaag 3420
ctaccaagac gaagaggatg aagaggatga ggaggcagat tgccttaat atattgacaa 3480
tactgataag ataatatatac ttttatatac aagatatcgc cgtatgtaaag gatttcaggg 3540
ggcaaggcat aggcagcgcg cttatcaata tatctataga atgggcaaag cataaaaact 3600
tgcatggact aatgcttgaa acccaggaca ataaccttat agcttgtaaa ttctatcata 3660
attgggtaat gactccaact tattgatagt gtttatgtt cagataatgc ccgatgactt 3720
tgtcatgcag ctccaccgat tttgagaacg acagcgactt ccgtcccagc cgtgccaggt 3780
gctgcctcag attcaggta tgccgctcaa ttgcgtcgt atatcgctt ctgattacgt 3840
gcagcttcc cttcaggcgg gattcataca gcccgcagcc atccgtcatc catatcacca 3900
cgtcaaaggg tgacagcagg ctcataagac gcccagcgt cgccatagtg cgttcaccga 3960
atacgtgcgc aacaaccgtc ttccggagac tgtcatacgc gtaaaacagc cagcgtggc 4020
gcgatttagc cccgacatag ccccactgtt cgtccatttc cgccgcacacg atgacgtcac 4080
tgccccggctg tatgcgcgag gttaccgact gcccgcgtg ttttttaagt gacgtaaaat 4140
cgtgttgggg ccaacgccc taatgcgggc tggtgcggg catccaacgc cattcatggc 4200
catatcaatg attttctggt gcgtaccggg ttgagaagcgt gtgtaaatgtc actgcagttg 4260
ccatgttta cggcagttag agcagagata gcccgcgtgatgt cggccgggtgc ttttgcgtt 4320
acgcaccacc ccgtcagtag ctgaacagga gggacagctg atagacacag aagccactgg 4380
agcacctcaa aaacaccatc atacactaaa tcagtaagtt ggcagcatca cccataattg 4440
tggtttcaaa atcggctccg tcgataactat gttatacgcc aactttgaaa acaactttga 4500
aaaagctgtt ttctggatt taaggttta gaatgcagg aacagtgaat tggagttcgt 4560

cttgttataa ttagcttctt gggtatctt taaaactgt agaaaagagg aaggaaataa 4620
 taaatggcta aaatgagaat atcaccggaa ttgaaaaaac tgatcgaaaa ataccgctgc 4680
 gtaaaagata cggaaggaat gtctcctgct aaggtatata agctggtggg agaaaatgaa 4740
 aacatatatt taaaatgac ggacagccgg tataaaggga ccacctatga tgtggaacgg 4800
 gaaaaggaca tgatgctatg gctgaaagga aagctgcctg ttccaaaggt cctgcacttt 4860
 gaacggcatg atggctggag caatctgctc atgagtgagg ccgatggcgt ccttgctcg 4920
 gaagagtatg aagatgaaca aagccctgaa aagattatcg agctgtatgc ggagtgcac 4980
 aggcttttc actccatcga catatcgat tgtccctata cgaatagctt agacagccgc 5040
 ttagccgaat tggattactt actgaataac gatctggccg atgtggattt cgaaaactgg 5100
 gaagaagaca ctccatttaa agatccgcgc gagctgtatg atttttaaa gacggaaaag 5160
 cccgaagagg aacttgtctt ttcccacggc gacctggag acagcaacat ctttgtaaa 5220
 gatggcaaag taagtggctt tattgatctt gggagaagcg gcagggcgga caagtggat 5280
 gacattgcct tctgcgtccg gtcgatcagg gaggatatcg gggagaaca gtagtcgag 5340
 ctatTTTtgcgacttacttacttggg gatcaagcctt gattgggaga aaataaaaata ttatTTTtaa 5400
 ctggatgaat tgTTTtagta cctagatgtg ggcgaacgt gcccggcaca agcaggagcg 5460
 cacccacttc ttccgcattca agtgtttgg ctctcaggcc gaggcccacg gcaagtattt 5520
 gggcaagggg tcgctggat tctgtcaggg caagattcg aataccaatg acgagaagga 5580
 cggccagacg gtctacggga ccgacttcat tgccgataag gtggattatc tggacaccaa 5640
 ggcaccaggc gggtaaaatc aggaataagg gcacattgcc ccggcgttag tcggggcaat 5700
 cccgcaagga gggtaatga atcgacgtt tgacoggaag gcatacaggc aagaactgat 5760
 cgacgcgggg tttccgcgg aggatgccga aaccatcgca agccgcaccc tcattgcgtgc 5820
 gccccgcgaa accttccagt ccgtcgctc gatggtccag caagctacgg ccaagatcga 5880
 ggcgcacaggc gtgcacactgg ctccccctgc cctgcccgcg ccattggccg ccgtggagcg 5940
 ttgcgtcgat ctcgaacagg aggccggcagg tttggcgaag tcgatgacca tcgacacgcg 6000
 aggaactatg acgaccaaga agcgaaaaac cgccggcgcg gacctggcaa aacaggtcag 6060
 cgaggccaaag caggccgcgt tgctgaaaca cacgaaggcag cagatcaagg aaatgcagct 6120
 ttccctgttc gatattgcgc cgtggccgga cacgatgcga gcgatgcac accacacggc 6180
 ccgctctgcc ctgttcacca cgcgcaacaa gaaaatcccg cgcgaggcgc tgcaaaacaa 6240
 ggtcattttc cacgtcaaca aggacgtgaa gatcacctac accggcgtcg agctgcgggc 6300

cgacgatgac gaactgggt gtgcagcagg gttggagtac gcgaagcgca cccctatcg 6360
 cgagccgatc accttcacgt tctacgagct ttgccaggac ctgggctggt cgatcaatgg 6420
 ccggtattac acgaaggccg aggaatgcct gtcgcgccta caggcgacgg cgatgggctt 6480
 cacgtccgac cgcggtggc acctggaatc ggtgtcgctg ctgcaccgct tccgcgtcct 6540
 ggaccgtggc aagaaaaacgt cccgttgcga ggtcctgatc gacgaggaaa tcgtcgtgct 6600
 gtttgctggc gaccactaca cgaattcat atgggagaag taccgcaagc tgtcgcccac 6660
 ggcccacgg atgttcgact atttcagctc gcaccggag cogtacccgc tcaagctgga 6720
 aacctccgc ctcatgtgcg gatcgattc caccgcgtg aagaagtggc gcgagcagg 6780
 cggcgaagcc tgcgaaagagt tgcgaggcag cggcctggt gaacacgcct gggtaatga 6840
 tgacctggtg cattgcaaacc gcttagggcct tgtgggtca gttccggctg gggttcagc 6900
 agccagcgct ttactggcat ttcaggaaca agcgggact gctcgacgca cttgcttcgc 6960
 tcagtatcgc tcgggacgca cggcgcgctc tacgaactgc cgataaacag aggattaaaa 7020
 ttgacaattt tgattaaggc tcagattcga cggcttggag cggccgacgt gcaggatttc 7080
 cgcgagatcc gattgtcggc cctgaagaaaa gctccagaga tttcgggtc cgtttacgag 7140
 cacgaggaga aaaagcccat ggaggcggtc gctgaacggt tgcgagatgc cgtggcattc 7200
 ggcgcctaca tcgacggcga gatcattggg ctgtcggtct tcaaacagga ggacggcccc 7260
 aaggacgctc acaaggcgca tctgtccggc gtttcgtgg agcccgaaca gcgaggccga 7320
 ggggtcgccg gtatgctgct gccccggttt cggcggtt tattgctcgat gatgatcgct 7380
 cgacagattc caacgggaat ctggtggtatc cgcatcttca tccctggcgactt 7440
 tcgctattct ggagcttggt gtttatttcg gtctaccgccc tgccggcggtt ggtcgccgg 7500
 acggtaggcg ctgtgcagcc gctgatggtc gtgttcatct ctggcgctct gctaggtac 7560
 ccgatacgat tgatggcggtt cctggggctt atttgcggaa ctgcggcggtt ggcgctgttg 7620
 gtgttgacac caaacgcagc gctagatcct gtcggcgctc cagcgggcctt ggcggggcg 7680
 gtttccatgg cggtcgaaac cgtgctgacc cgcaagtggc aacctcccgat gcctctgctc 7740
 accttacccg cctggcaact ggcggccggaa ggacttctgc tggccatggc ggtcgccgg 7800
 tttgatccgc caatcccgat gcctacagga accaatgttc tggccatggc ggtcgccgg 7860
 ctgatcgag cgggtttaaac ctacttcctt tggccatggc ggtcgccgg actcgaaacct 7920
 acagttgttt ctttactggg ctttctcagc cccagatctg ggtcgatca gcccggatg 7980
 catcaggccg acagtcggaa ctgcgggtcc ccgacctgtt ccattcggtt agcaatggat 8040

aggggagttg atatcgtaa cgttcacttc taaagaata ggcactca gcttcctcag 8100
 cggcttatac cagcgatttc ctattatgtc ggcatagttc tcaagatcga cagoctgtca 8160
 cggtaagcg agaaatgaat aagaaggctg ataattcgg a tctctgcgag ggagatgata 8220
 tttgatcaca ggcagcaacg ctctgtcatc gttacaatca acatgctacc ctccgcgaga 8280
 tcataccgtgt ttcaaaccgg gcagcttagt tgccgttctt ccgaatagca tcggtaacat 8340
 gagcaaagtc tgccgcctta caacggctct cccgctgacg ccgtcccgga ctgatggct 8400
 gcctgtatcg agtggtgatt ttgtggcag ctgcccgtcg gggagctgtt ggctggctgg 8460
 tggcaggata tattgtggtg taaacaaatt gacgcttaga caacttaata acacattgcg 8520
 gacgtttta atgtactggg gtggtttttc ttttaccagg tgagacggc aacagctgat 8580
 tgcccttcac cgccctggccc tgagagagtt gcagcaagcg gtccacgctg gttgccccca 8640
 gcaggcgaaa atcctgtttt atggtggttc cggaaatcgcc aaaatccctt ataaatcaaa 8700
 agaatagccc gagatagggt tgagtgttgt tccagttgg aacaagatc cactattaaa 8760
 gaacgtggac tccaacgtca aaggcgaaa aaccgtctat cagggcgatg gcccactacg 8820
 tgaaccatca cccaaatcaa gtttttggg gtcgaggtgc cgtaaagcac taaatcgaa 8880
 ccctaaaggg agccccccgat ttagagctt acggggaaag cggcgaaacg tggcgagaaa 8940
 ggaagggaag aaagcgaaag gagcggcgc cattcaggct ggcactgt tgggaaggc 9000
 gatcggtgcg ggcctctcg ctattacgaa agctggcgaa agggggatgt gctgcaaggc 9060
 gattaagttt ggtAACGCCA gggtttccc agtcacgacg ttgtaaaacg acggccagt 9120
 aattaattcc catcttgaaa gaaatatagt taaaatattt attgataaaaa taacaagtca 9180
 ggtattatag tccaagcaaa aacataaatt tattgatgca agttttaaatt cagaaatatt 9240
 tcaataactg attatatcag ctggtagatt gccgtatgt aaagactgag tgctgatatta 9300
 tgtgtataac ataaattgtat gatatagtca gcttagctca tcgggggatc cgctgaaact 9360
 agcttgggtc ccgctcagaa gaactcgta agaaggcgat agaaggcgat ggcgtcgaa 9420
 tcgggagcgg cgataccgta aagcacgagg aagcggtcag cccattcgcc gccaagctct 9480
 tcagcaatat cacgggtac caacgctatg tcctgtatgc ggtccgccc acccagccgg 9540
 ccacagtcga tgaatccaga aaagcgccca tttccacca tgatattcgg caagcaggca 9600
 tcgccatggg tcacgacgag atcctcgccg tcgggcatgc ggccttgag cctggcgaac 9660
 agttcggctg ggcgcgagccc ctgatgtctc tgcgtccagat catcctgatc gacaagaccg 9720
 gcttccatcc gagtaacgtgc tcgctcgatg cgatgtttcg ctgggtggtc gaatggcag 9780

gtagccggat caagcgtatg cagccgccgc attgcatcg ccatgatgga tactttctcg 9840
 gcaggagcaa ggtgagatga caggagatcc tgccccggca ctgcgccaa tagcagccag 9900
 tcccttcccgc ttcaagtgc aacgtcgagc acagctgcgc aaggaacgcc cgtcgtggcc 9960
 agccacgata gccgcgcgtgc ctgcgtcctgc agttcattca gggcacccgga caggtcggtc 10020
 ttgacaaaaaa gaaccgggcg cccctgcgc gacagccgga acacggccgc atcagagcag 10080
 ccgatttgtct gttgtgcaca gtcatagccg aatagcctct ccacccaagc ggccggagaa 10140
 cctgcgtgca atccatcttg ttcaatccaa gtcacccatgg gccctcgact agagtcgaga 10200
 tctggattga gägtgaatat gagactctaa ttggataccg aggggaattt atggaacggtc 10260
 agtggagcat ttttgcacaag aaatatttgc tagctgatag tgacctttagg cgacttttga 10320
 acgcgcacaata atggtttctg acgtatgtc tttagtcatt aaactccaga aaccgcggc 10380
 tgagtggctc cttcaacggtt gcggttctgt cagttccaaa cgtaaaacgg ctgtccgc 10440
 gtcatcgccg ggggtcataa cgtgactccc ttaattctcc gctcatgatc ttgatccct 10500
 gcccgcacatcag atccttggcg gcaagaaaagc catccagttt actttgcagg gcttcccaac 10560
 cttaccagag ggcgcggccag ctggcaattc cggttcgctt gctgtccata aaaccgcggc 10620
 gtcttagctat cgccatgtaa gcccactgca agctacactgc tttctcttgc cgcttgcgtt 10680
 ttcccttgc cagatagccc agtagctgac attcatccgg ggtcagoacc gtttctgcgg 10740
 actggcttac tacgtgttcc gtttcctta gcagcccttg cgccctgagt gcttgeggca 10800
 gcgtaagct tgcatgcctg caggtcgacg gcgccggag ctccctcgagc aaatttacac 10860
 attgccacta aacgtctaaa cccttgcata ttgttttgc tttactatgt gtgttatgta 10920
 tttgatttgc gataaaatttt tatatttggc actaaattta taacaccctt tatgctaacg 10980
 tttgccaaca ctttagcaatt tgcaagttga ttaatttgcatt ctaaatttattt tttgtcttct 11040
 aaatacatat actaatcaac tggaaatgta aatatttgc aatatttcta ctataggaga 11100
 attaaagtga gtgaatatgg taccacaagg tttggagatt taatttgc aatgctgcatt 11160
 ggtatggcata tacacccaaac attcaataat tcttgaggat aataatggta ccacacaaga 11220
 tttgaggtgc atgaacgtca cgtggacaaa aggttttagta atttttcaag acaacaatgt 11280
 taccacacac aagtttttagt gtgcattgcatt ggtatggccctg tggaaagttt aaaaatattt 11340
 tggaaatgat ttgcattggaa gccatgtgta aaaccatgac atccacttgg aggatgcaat 11400
 aatgaagaaa actacaaatt tacatgcaac tagttatgca ttttttttatataatgagga 11460
 ttttgcata ctttcattca tacacactca ctaagttta cacgattata atttcttcat 11520

agccagccca ccgcgggtgga aa atg gag gtc gtg gag aga ttc tac ggt gag 11572
 Met Glu Val Val Glu Arg Phe Tyr Gly Glu
 1 5 10

ttg gat ggg aag gtc tcg cag ggc gtg aat gca ttg ctg ggt agt ttt 11620
 Leu Asp Gly Lys Val Ser Gln Gly Val Asn Ala Leu Leu Gly Ser Phe
 15 20 25

ggg gtg gag ttg acg gat acg ccc act acc aaa ggc ttg ccc ctc gtt 11668
 Gly Val Glu Leu Thr Asp Thr Pro Thr Lys Gly Leu Pro Leu Val
 30 35 40

gac agt ccc aca ccc atc gtc ctc ggt gtt tct gta tac ttg act att 11716
 Asp Ser Pro Thr Pro Ile Val Leu Gly Val Ser Val Tyr Leu Thr Ile
 45 50 55

gtc att gga ggg ctt ttg tgg ata aag gcc agg gat ctg aaa ccg cgc 11764
 Val Ile Gly Gly Leu Leu Trp Ile Lys Ala Arg Asp Leu Lys Pro Arg
 60 65 70

gcc tcg gag cca ttt ttg ctc caa gct ttg gtg ctt gtg cac aac ctg 11812
 Ala Ser Glu Pro Phe Leu Leu Gln Ala Leu Val Leu Val His Asn Leu
 75 80 85 90

ttc tgt ttt gcg ctc agt ctg tat atg tgc gtg ggc atc gct tat cag 11860
 Phe Cys Phe Ala Leu Ser Leu Tyr Met Cys Val Gly Ile Ala Tyr Gln
 95 100 105

gct att acc tgg cgg tac tct ctc tgg ggc aat gca tac aat cct aaa 11908
 Ala Ile Thr Trp Arg Tyr Ser Leu Trp Gly Asn Ala Tyr Asn Pro Lys
 110 115 120

cat aaa gag atg gcg att ctg gta tac ttg ttc tac atg tct aag tac 11956
 His Lys Glu Met Ala Ile Leu Val Tyr Leu Phe Tyr Met Ser Lys Tyr
 125 130 135

gtg gaa ttc atg gat acc gtt atc atg ata ctg aag cgc agc acc agg 12004
 Val Glu Phe Met Asp Thr Val Ile Met Ile Leu Lys Arg Ser Thr Arg
 140 145 150

caa ata agc ttc ctc cac gtt tat cat cat tct tca att tcc ctc att 12052
 Gln Ile Ser Phe Leu His Val Tyr His Ser Ser Ile Ser Leu Ile
 155 160 165 170

tgg tgg gct att gct cat cac gct cct ggc ggt gaa gca tat tgg tct 12100
 Trp Trp Ala Ile Ala His His Ala Pro Gly Gly Glu Ala Tyr Trp Ser
 175 180 185

gcg gct ctg aac tca gga gtg cat gtt ctc atg tat gcg tat tac ttc 12148
 Ala Ala Leu Asn Ser Gly Val His Val Leu Met Tyr Ala Tyr Tyr Phe
 190 195 200

ttg gct gcc tgc ctt cga agt agc cca aag tta aaa aat aag tac ctt 12196
 Leu Ala Ala Cys Leu Arg Ser Ser Pro Lys Leu Lys Asn Lys Tyr Leu
 205 210 215

ttt tgg ggc agg tac ttg aca caa ttc caa atg ttc cag ttt atg ctg 12244
 Phe Trp Gly Arg Tyr Leu Thr Gln Phe Gln Met Phe Gln Phe Met Leu

220	225	230	
aac tta gtg cag gct tac tac gac atg aaa acg aat gcg cca tat cca Asn Leu Val Gln Ala Tyr Tyr Asp Met Lys Thr Asn Ala Pro Tyr Pro 235 240 245 250 12292			
caa tgg ctg atc aag att ttg ttc tac tac atg atc tcg ttg ctg ttt Gln Trp Leu Ile Lys Ile Leu Phe Tyr Tyr Met Ile Ser Leu Leu Phe 255 260 265 12340			
ctt ttc ggc aat ttt tac gta caa aaa tac atc aaa ccc tct gac gga Leu Phe Gly Asn Phe Tyr Val Gln Lys Tyr Ile Lys Pro Ser Asp Gly 270 275 280 12388			
aag caa aag gga gct aaa act gag tga tctagaaggc ctcctgcttt Lys Gln Lys Gly Ala Lys Thr Glu 285 290 12435			
aatgagatat gcgagacgccc tatgatecgca tgatatttgc tttcaattct gttgtgcacg ttgtaaaaaaaaa cctgagcatg tgttagctcgat atccttaccgc cgggtttcgg ttcattctaa 12555			12495
tgaatatatac acccgttact atcgtatTTT tatgaataat attctccgtt caatttactg attgtccgtc gagcaaattt acacattgcc actaaacgtc taaacccttg taatttgttt 12615			
ttgtttact atgtgtgtta tgtatttgat ttgcgataaaa ttttatatt tggactaaa 12735			
tttataacac ctttatgct aacgttgcc aacacttagc aatttgcag ttgattaatt 12795			
gattctaaat tattttgtc ttctaaatac atatactaattt caactggaaa tgtaaatatt 12855			
tgctaataattt tctactatag gagaattaaa gtgagtgaat atggtaccac aaggtttgga 12915			
gatttaatttggatgg catatacacc aaacattcaa taattcttga 12975			
ggataataat ggtaccacac aagatttgag gtgcgtgaac gtcacgtgga caaaagggttt 13035			
agtaattttt caagacaaca atgttaccac acacaagttt tgaggtgtcat gcatggatgc 13095			
cctgtggaaa gttaaaaat attttggaaa tgatttgcattt ggaagccatg tgtaaaacca 13155			
tgacatccac ttggaggatg caataatgaa gaaaactaca aatttacatg caactagtta 13215			
tgcattgtatg ctatataatg aggatttgc aatactttca ttcatacaca ctcactaagt 13275			
tttacacgtatataatttct tcatagccag cggatcc atg gta ttc gct ggc ggt Met Val Phe Ala Gly Gly 295 13330			
gga ctt cag cag ggc tct ctc gaa gaa aac atc gac gtc gag cac att Gly Leu Gln Gln Gly Ser Leu Glu Glu Asn Ile Asp Val Glu His Ile 300 305 310 13378			
gcc agt atg tct ctc ttc agc gac ttc ttc agt tat gtg tct tca act Ala Ser Met Ser Leu Phe Ser Asp Phe Phe Ser Tyr Val Ser Ser Thr 315 320 325 13426			

100

gtt ggt tcg tgg agc gta cac agt ata caa cct ttg aag cgc ctg acg 13474
 Val Gly Ser Trp Ser Val His Ser Ile Gln Pro Leu Lys Arg Leu Thr
 330 335 340 345

agt aag aag cgt gtt tcg gaa agc gct gcc gtg caa tgt ata tca gct 13522
 Ser Lys Lys Arg Val Ser Ala Ala Val Gln Cys Ile Ser Ala
 350 355 360

gaa gtt cag aga aat tcg agt acc cag gga act gcg gag gca ctc gca 13570
 Glu Val Gln Arg Asn Ser Ser Thr Gln Gly Thr Ala Glu Ala Leu Ala
 365 370 375

gaa tca gtc gtg aag ccc acg aga cga agg tca tct cag tgg aag aag 13618
 Glu Ser Val Val Lys Pro Thr Arg Arg Ser Ser Gln Trp Lys Lys
 380 385 390

tcg aca cac ccc cta tca gaa gta gca gta cac aac aag cca agc gat 13666
 Ser Thr His Pro Leu Ser Glu Val Ala Val His Asn Lys Pro Ser Asp
 395 400 405

tgc tgg att gtt gta aaa aac aag gtg tat gat gtt tcc aat ttt gcg 13714
 Cys Trp Ile Val Val Lys Asn Lys Val Tyr Asp Val Ser Asn Phe Ala
 410 415 420 425

gac gag cat ccc gga gga tca gtt att agt act tat ttt gga cga gac 13762
 Asp Glu His Pro Gly Gly Ser Val Ile Ser Thr Tyr Phe Gly Arg Asp
 430 435 440

ggc aca gat gtt ttc tct agt ttt cat gca gct tct aca tgg aaa att 13810
 Gly Thr Asp Val Phe Ser Ser Phe His Ala Ala Ser Thr Trp Lys Ile
 445 450 455

ctt caa gac ttt tac att ggt gac gtg gag agg gtg gag ccg act cca 13858
 Leu Gln Asp Phe Tyr Ile Gly Asp Val Glu Arg Val Glu Pro Thr Pro
 460 465 470

gag ctg ctg aaa gat ttc cga gaa atg aga gct ctt ttc ctg agg gag 13906
 Glu Leu Leu Lys Asp Phe Arg Glu Met Arg Ala Leu Phe Leu Arg Glu
 475 480 485

caa ctt ttc aaa agt tcg aaa ttg tac tat gtt atg aag ctg ctc acg 13954
 Gln Leu Phe Lys Ser Ser Lys Leu Tyr Tyr Val Met Lys Leu Leu Thr
 490 495 500 505

aat gtt gct att ttt gct gcg agc att gca ata ata tgt tgg agc aag 14002
 Asn Val Ala Ile Phe Ala Ala Ser Ile Ala Ile Ile Cys Trp Ser Lys
 510 515 520

act att tca gcg gtt ttg gct tca gct tgt atg atg gct ctg tgt ttc 14050
 Thr Ile Ser Ala Val Leu Ala Ser Ala Cys Met Met Ala Leu Cys Phe
 525 530 535

caa cag tgc gga tgg cta tcc cat gat ttt ctc cac aat cag gtg ttt 14098
 Gln Gln Cys Gly Trp Leu Ser His Asp Phe Leu His Asn Gln Val Phe
 540 545 550

gag aca cgc tgg ctt aat gaa gtt gtc ggg tat gtg atc ggc aac gcc 14146
 Glu Thr Arg Trp Leu Asn Glu Val Val Gly Tyr Val Ile Gly Asn Ala

101

555	560	565	
gtt ctg ggg ttt agt aca ggg tgg tgg aag gag aag cat aac ctt cat Val Leu Gly Phe Ser Thr Gly Trp Trp Lys Glu Lys His Asn Leu His 570 575 580 585 14194			
cat gct gct cca aat gaa tgc gat cag act tac caa cca att gat gaa His Ala Ala Pro Asn Glu Cys Asp Gln Thr Tyr Gln Pro Ile Asp Glu 590 595 600 14242			
gat att gat act ctc ccc ctc att gcc tgg agc aag gac ata ctg gcc Asp Ile Asp Thr Leu Pro Leu Ile Ala Trp Ser Lys Asp Ile Leu Ala 605 610 615 14290			
aca gtt gag aat aag aca ttc ttg cga atc ctc caa tac cag cat ctg Thr Val Glu Asn Lys Thr Phe Leu Arg Ile Leu Gln Tyr Gln His Leu 620 625 630 14338			
ttc ttc atg ggt ctg tta ttt ttc gcc cgt ggt agt tgg ctc ttt tgg Phe Phe Met Gly Leu Leu Phe Phe Ala Arg Gly Ser Trp Leu Phe Trp 635 640 645 14386			
agc tgg aga tat acc tct aca gca gtg ctc tca cct gtc gac agg ttg Ser Trp Arg Tyr Thr Ser Thr Ala Val Leu Ser Pro Val Asp Arg Leu 650 655 660 665 14434			
ttg gag aag gga act gtt ctg ttt cac tac ttt tgg ttc gtc ggg aca Leu Glu Lys Gly Thr Val Leu Phe His Tyr Phe Trp Phe Val Gly Thr 670 675 680 14482			
gcg tgc tat ctt ctc cct ggt tgg aag cca tta gta tgg atg gcg gtg Ala Cys Tyr Leu Leu Pro Gly Trp Lys Pro Leu Val Trp Met Ala Val 685 690 695 14530			
act gag ctc atg tcc ggc atg ctg ctg ggc ttt gta ttt gta ctt agc Thr Glu Leu Met Ser Gly Met Leu Leu Gly Phe Val Phe Val Leu Ser 700 705 710 14578			
cac aat ggg atg gag gtt tat aat tcg tct aaa gaa ttc gtg agt gca His Asn Gly Met Glu Val Tyr Asn Ser Ser Lys Glu Phe Val Ser Ala 715 720 725 14626			
cag atc gta tcc aca cgg gat atc aaa gga aac ata ttc aac gac tgg Gln Ile Val Ser Thr Arg Asp Ile Lys Gly Asn Ile Phe Asn Asp Trp 730 735 740 745 14674			
ttc act ggt ggc ctt aac agg caa ata gag cat cat ctt ttc cca aca Phe Thr Gly Gly Leu Asn Arg Gln Ile Glu His His Leu Phe Pro Thr 750 755 760 14722			
atg ccc agg cat aat tta aac aaa ata gca cct aga gtg gag gtg ttc Met Pro Arg His Asn Leu Asn Lys Ile Ala Pro Arg Val Glu Val Phe 765 770 775 14770			
tgt aag aaa caç ggt ctg gtg tac gaa gac gta tct att gct acc ggc Cys Lys Lys His Gly Leu Val Tyr Glu Asp Val Ser Ile Ala Thr Gly 780 785 790 14818			

act tgc aag gtt ttg aaa gca ttg aag gaa gtc gcg gag gct gcg gca 14866
 Thr Cys Lys Val Leu Lys Ala Leu Lys Glu Val Ala Glu Ala Ala Ala
 795 800 805

gag cag cat gct acc acc agt taa gctacgtta accctgctt aatgagat 14920
 Glu Gln His Ala Thr Thr Ser
 810 815

gcgagacgcc tatgatcgca tgatattgc ttcaattct gttgtgcacg ttgtaaaaaa 14980
 cctgagcatg ttagctcag atcccttaccg ccggttcgg ttcattctaa tgaatatatc 15040
 acccgttact atcgatattt tatgaataat attctccgtt caatttactg attgtccgtc 15100
 gagcaaattt acacattgcc actaaacgtc taaacccttg taatttgtt ttgtttact 15160
 atgtgtgtta tgtatttgat ttgcgataaaa ttttatatt tggtactaaa tttataacac 15220
 ctttatgct aacgtttgcc aacacttagc aatttgcag ttgattaatt gattctaaat 15280
 tattttgtc ttctaaatac atatacta caactggaaa tgtaaatatt tgctaatatt 15340
 tctactatag gagaattaaa gtgagtgaat atggtaccac aaggtttggaa gatthaattg 15400
 ttgcaatgct gcatggatgg catatacacc aaacattcaa taattctga ggataataat 15460
 ggtaccacac aagatttgag gtgcgtgaac gtcacgtgga caaaaggttt agtaatttt 15520
 caagacaaca atgttaccac acacaagttt tgaggtgcat gcatggatgc cctgtggaaa 15580
 gttaaaaat attttggaaa tgatttgcat ggaagccatg tgtaaaacca tgacatccac 15640
 ttggaggatg caataatgaa gaaaactaca aatttacatg caactagtt tgcatgttgt 15700
 ctatataatg aggatttgc aatactttca ttcatcacaca ctcactaagt tttacacgat 15760
 tataattct tcataccag cagatctaaa atg gct ccg gat ggc gat aag ctt 15814
 Met Ala Pro Asp Ala Asp Lys Leu
 820 825

cga caa cgc cag acg act gcg gta gcg aag cac aat gct gct acc ata 15862
 Arg Gln Arg Gln Thr Thr Ala Val Ala Lys His Asn Ala Ala Thr Ile
 830 835 840

tcg acg cag gaa cgc ctt tgc agt ctg tct tcg ctc aaa ggc gaa gaa 15910
 Ser Thr Gln Glu Arg Leu Cys Ser Leu Ser Ser Leu Lys Gly Glu Glu
 845 850 855

gtc tgc atc gac gga atc atc tat gac ctc caa tca ttc gat cat ccc 15958
 Val Cys Ile Asp Gly Ile Ile Tyr Asp Leu Gln Ser Phe Asp His Pro
 860 865 870

ggg ggt gaa acg atc aaa atg ttt ggt ggc aac gat gtc act gta cag 16006
 Gly Gly Glu Thr Ile Lys Met Phe Gly Gly Asn Asp Val Thr Val Gln
 875 880 885

tac aag atg att cac ccg tac cat acc gag aag cat ttg gaa aag atg 16054
 Tyr Lys Met Ile His Pro Tyr His Thr Glu Lys His Leu Glu Lys Met

103

890	895	900	905	
aag cgt gtc ggc aag gtg acg gat ttc gtc tgc gag tac aag ttc gat Lys Arg Val Gly Lys Val Thr Asp Phe Val Cys Glu Tyr Lys Phe Asp 910		915		920
acc gaa ttt gaa cgc gaa atc aaa cga gaa gtc ttc aag att gtg cga Thr Glu Phe Glu Arg Glu Ile Lys Arg Glu Val Phe Lys Ile Val Arg 925		930		935
cga ggc aag gat ttc ggt act ttg gga tgg ttc ttc cgt gcg ttt tgc Arg Gly Lys Asp Phe Gly Thr Leu Gly Trp Phe Phe Arg Ala Phe Cys 940		945		950
tac att gcc att ttc ttc tac ctg cag tac cat tgg gtc acc acc gga Tyr Ile Ala Ile Phe Phe Tyr Leu Gln Tyr His Trp Val Thr Thr Gly 955		960		965
acc tct tgg ctg ctg gcc gtg gcc tac gga atc tcc caa gcg atg att Thr Ser Trp Leu Leu Ala Val Ala Tyr Gly Ile Ser Gln Ala Met Ile 970		975		985
ggc atg aat gtc cag cac gat gcc aac cac ggg gcc acc tcc aag cgt Gly Met Asn Val Gln His Asp Ala Asn His Gly Ala Thr Ser Lys Arg 990		995		1000
ccc tgg gtc aac gac atg cta ggc ctc ggt gcg gat ttt att ggt ggt Pro Trp Val Asn Asp Met Leu Gly Leu Gly Ala Asp Phe Ile Gly Gly 1005		1010		1015
tcc aag tgg ctc tgg cag gaa caa cac tgg acc cac cac gct tac acc Ser Lys Trp Leu Trp Gln Glu Gln His Trp Thr His His Ala Tyr Thr 1020		1025		1030
aat cac gcc gag atg ccc gat agc ttt ggt gcc gaa cca atg ctc Asn His Ala Glu Met Asp Pro Asp Ser Phe Gly Ala Glu Pro Met Leu 1035		1040		1045
cta ttc aac gac tat ccc ttg gat cat ccc gct cgt acc tgg cta cat Leu Phe Asn Asp Tyr Pro Leu Asp His Pro Ala Arg Thr Trp Leu His 1050		1055		1060
cgc ttt caa gca ttc ttt tac atg ccc gtc ttg gct gga tac tgg ttg Arg Phe Gln Ala Phe Phe Tyr Met Pro Val Leu Ala Gly Tyr Trp Leu 1070		1075		1080
tcc gct gtc ttc aat cca caa att ctt gac ctc cag caa cgc ggc gca Ser Ala Val Phe Asn Pro Gln Ile Leu Asp Leu Gln Gln Arg Gly Ala 1085		1090		1095
ctt tcc gtc ggt atc cgt ctc gac aac gct ttc att cac tcg cga cgc Leu Ser Val Gly Ile Arg Leu Asp Asn Ala Phe Ile His Ser Arg Arg 1100		1105		1110
aag tat gcg gtt ttc tgg cgg gct gtg tac att gcg gtg aac gtg att Lys Tyr Ala Val Phe Trp Arg Ala Val Tyr Ile Ala Val Asn Val Ile 1115		1120		1125

gct ccg ttt tac aca aac tcc ggc ctc gaa tgg tcc tgg cgt gtc ttt 16774
 Ala Pro Phe Tyr Thr Asn Ser Gly Leu Glu Trp Ser Trp Arg Val Phe
 1130 1135 1140 1145

 gga aac atc atg ctc atg ggt gtg gcg gaa tcg ctc gcg ctg gcg gtc 16822
 Gly Asn Ile Met Leu Met Gly Val Ala Glu Ser Leu Ala Leu Ala Val
 1150 1155 1160

 ctg ttt tcg ttg tcg cac aat ttc gaa tcc gcg gat cgc gat ccg acc 16870
 Leu Phe Ser Leu Ser His Asn Phe Glu Ser Ala Asp Arg Asp Pro Thr
 1165 1170 1175

 gcc cca ctg aaa aag acg gga gaa cca gtc gac tgg ttc aag aca cag 16918
 Ala Pro Leu Lys Lys Thr Gly Glu Pro Val Asp Trp Phe Lys Thr Gln
 1180 1185 1190

 gtc gaa act tcc tgc act tac ggt gga ttc ctt tcc ggt tgc ttc acg 16966
 Val Glu Thr Ser Cys Thr Tyr Gly Phe Leu Ser Gly Cys Phe Thr
 1195 1200 1205

 gga ggt ctc aac ttt cag gtt gaa cac cac ttg ttc cca cgc atg agc 17014
 Gly Gly Leu Asn Phe Gln Val Glu His His Leu Phe Pro Arg Met Ser
 1210 1215 1220 1225

 agc gct tgg tat ccc tac att gcc ccc aag gtc cgc gaa att tgc gcc 17062
 Ser Ala Trp Tyr Pro Tyr Ile Ala Pro Lys Val Arg Glu Ile Cys Ala
 1230 1235 1240

 aaa cac ggc gtc cac tac gac tac tac ccg tgg atc cac caa aac ttt 17110
 Lys His Gly Val His Tyr Ala Tyr Tyr Pro Trp Ile His Gln Asn Phe
 1245 1250 1255

 ctc tcc acc gtc cgc tac atg cac gcg gcc ggg acc ggt gcc aac tgg 17158
 Leu Ser Thr Val Arg Tyr Met His Ala Ala Gly Thr Gly Ala Asn Trp
 1260 1265 1270

 cgc cag atg gcc aga gaa aat ccc ttg acc gga cgg gcg taa 17200
 Arg Gln Met Ala Arg Glu Asn Pro Leu Thr Gly Arg Ala
 1275 1280 1285

 agatctgccg gcatcgatcc cggccatgg cctgctttaa tgagatatgc gagacgccta 17260
 tgatcgatg atatttgctt tcaattctgt tgtgcacgtt gtaaaaaacc tgagcatgtg 17320
 tagctcagat cttaccgccc gtttcgggtt cattctaattg aatataatcac ccgttactat 17380
 cgtatttta tgaataaatat tctccgttca atttactgat tgtccgtcga cgagctcggc 17440
 ggcctcttag aggatcgatg aattcagatc ggctgagtgg ctccctcaac gttgcgggttc 17500
 tgtcagttcc aaacgtaaaa cggcttgtcc cgcgtcatcg gccccgggtca taacgtgact 17560
 cccttaattc tccgctcatg atcagattgt cgtttccgc cttcagttta aactatcagt 17620
 gtttgacagg atatattggc gggtaaacct aagagaaaag agcgtttatt agaataatcg 17680
 gatatttaaa agggcgtgaa aaggtttatac cttcgtccat ttgtatgtgc atgccaaccca 17740

cagggttcccc ca

17752

<210> 29
<211> 290
<212> PRT
<213> Unknown

<400> 29

Met Glu Val Val Glu Arg Phe Tyr Gly Gly Glu Leu Asp Gly Lys Val Ser
 1 5 10 15

Val Leu Gly Val Ser Val Tyr Leu Thr Ile Val Ile Gly Gly Leu Leu
50 55 60

Leu Gln Ala Leu Val Leu Val His Asn Leu Phe Cys Phe Ala Leu Ser
85 90 95

Leu Tyr Met Cys Val Gly Ile Ala Tyr Gln Ala Ile Thr Trp Arg Tyr
100 105 110

Ser Leu Trp Gly Asn Ala Tyr Asn Pro Lys His Lys Glu Met Ala Ile
115 120 125

Leu Val Tyr Leu Phe Tyr Met Ser Lys Tyr Val Glu Phe Met Asp Thr
 130 135 140

Val Ile Met Ile Leu Lys Arg Ser Thr Arg Gin Ile Ser Phe Leu His
145 150 155 160

Val Tyr His His Ser Ser Ile Ser Leu Ile Trp Trp Ala Ala Ile Ala His
 165 170 175

His Ala Pro Gly Gly Glu Ala Tyr Trp Ser Ala Ala Leu Asn Ser Gly
180 185 190

Val His Val Leu Met Tyr Ala Tyr Tyr Phe Leu Ala Ala Cys Leu Arg
195 200 205

Ser Ser Pro Lys Leu Lys Asn Lys Tyr Leu Phe Trp Gly Arg Tyr Leu
210 . 215 220

Thr Gln Phe Gln Met Phe Gln Phe Met Leu Asn Leu Val Gln Ala Tyr
225 . . . 230 . . . 235 . . . 240

Tyr Asp Met Lys Thr Asn Ala Pro Tyr Pro Gln Trp Leu Ile Lys Ile
245 250 255

Leu Phe Tyr Tyr Met Ile Ser Leu Leu Phe Leu Phe Gly Asn Phe Tyr

106

260 265 270

Val Gln Lys Tyr Ile Lys Pro Ser Asp Gly Lys Gln Lys Gly Ala Lys
275 280 285

Thr Glu
290

<210> 30
<211> 525
<212> PRT
<213> Unknown

<400> 30
Met Val Phe Ala Gly Gly Leu Gln Gln Gly Ser Leu Glu Glu Asn
1 5 . 10 15

Ile Asp Val Glu His Ile Ala Ser Met Ser Leu Phe Ser Asp Phe Phe
20 25 30

Ser Tyr Val Ser Ser Thr Val Gly Ser Trp Ser Val His Ser Ile Gln
35 40 45

Pro Leu Lys Arg Leu Thr Ser Lys Lys Arg Val Ser Glu Ser Ala Ala
50 55 60

Ser Ser Gln Trp Lys Lys Ser Thr His Pro Leu Ser Glu Val Ala Val
100 105 110

His Asn Lys Pro Ser Asp Cys Trp Ile Val Val Lys Asn Lys Val Tyr
115 . 120 . 125

Asp Val Ser Asn Phe Ala Asp Glu His Pro Gly Gly Ser Val Ile Ser
130 135 140

Thr Tyr Phe Gly Arg Asp Gly Thr Asp Val Phe Ser Ser Phe His Ala
145 150 155 160

Ala Ser Thr Trp Lys Ile Leu Gln Asp Phe Tyr Ile Gly Asp Val Glu
165 170 175

Arg Val Glu Pro Thr Pro Glu Leu Leu Lys Asp Phe Arg Glu Met Arg
180 185 190

Ala Leu Phe Leu Arg Glu Gln Leu Phe Lys Ser Ser Lys Leu Tyr Tyr
195 200 205

Val Met Lys Leu Leu Thr Asn Val Ala Ile Phe Ala Ala Ser Ile Ala
210 215 220

Ile Ile Cys Trp Ser Lys Thr Ile Ser Ala Val Leu Ala Ser Ala Cys

225	230	235	240
Met Met Ala Leu Cys Phe Gln Gln Cys Gly Trp Leu Ser His Asp Phe			
245	250	255	
Leu His Asn Gln Val Phe Glu Thr Arg Trp Leu Asn Glu Val Val Gly			
260	265	270	
Tyr Val Ile Gly Asn Ala Val Leu Gly Phe Ser Thr Gly Trp Trp Lys			
275	280	285	
Glu Lys His Asn Leu His His Ala Ala Pro Asn Glu Cys Asp Gln Thr			
290	295	300	
Tyr Gln Pro Ile Asp Glu Asp Ile Asp Thr Leu Pro Leu Ile Ala Trp			
305	310	315	320
Ser Lys Asp Ile Leu Ala Thr Val Glu Asn Lys Thr Phe Leu Arg Ile			
325	330	335	
Leu Gln Tyr Gln His Leu Phe Phe Met Gly Leu Leu Phe Phe Ala Arg			
340	345	350	
Gly Ser Trp Leu Phe Trp Ser Trp Arg Tyr Thr Ser Thr Ala Val Leu			
355	360	365	
Ser Pro Val Asp Arg Leu Leu Glu Lys Gly Thr Val Leu Phe His Tyr			
370	375	380	
Phe Trp Phe Val Gly Thr Ala Cys Tyr Leu Leu Pro Gly Trp Lys Pro			
385	390	395	400
Leu Val Trp Met Ala Val Thr Glu Leu Met Ser Gly Met Leu Leu Gly			
405	410	415	
Phe Val Phe Val Leu Ser His Asn Gly Met Glu Val Tyr Asn Ser Ser			
420	425	430	
Lys Glu Phe Val Ser Ala Gln Ile Val Ser Thr Arg Asp Ile Lys Gly			
435	440	445	
Asn Ile Phe Asn Asp Trp Phe Thr Gly Gly Leu Asn Arg Gln Ile Glu			
450	455	460	
His His Leu Phe Pro Thr Met Pro Arg His Asn Leu Asn Lys Ile Ala			
465	470	475	480
Pro Arg Val Glu Val Phe Cys Lys Lys His Gly Leu Val Tyr Glu Asp			
485	490	495	
Val Ser Ile Ala Thr Gly Thr Cys Lys Val Leu Lys Ala Leu Lys Glu			
500	505	510	
Val Ala Glu Ala Ala Ala Glu Gln His Ala Thr Thr Ser			
515	520	525	

<211> 469

<212> PRT

<213> Unknown

<400> 31

Met	Ala	Pro	Asp	Ala	Asp	Lys	Leu	Arg	Gln	Arg	Gln	Thr	Thr	Ala	Val
1				5				10					15		

Ala	Lys	His	Asn	Ala	Ala	Thr	Ile	Ser	Thr	Gln	Glu	Arg	Leu	Cys	Ser
							20		25			30			

Leu	Ser	Ser	Leu	Lys	Gly	Glu	Glu	Val	Cys	Ile	Asp	Gly	Ile	Ile	Tyr
							35		40			45			

Asp	Leu	Gln	Ser	Phe	Asp	His	Pro	Gly	Gly	Glu	Thr	Ile	Lys	Met	Phe
							50		55			60			

Gly	Gly	Asn	Asp	Val	Thr	Val	Gln	Tyr	Lys	Met	Ile	His	Pro	Tyr	His
							65		70		75		80		

Thr	Glu	Lys	His	Leu	Glu	Lys	Met	Lys	Arg	Val	Gly	Lys	Val	Thr	Asp
							85		90			95			

Phe	Val	Cys	Glu	Tyr	Lys	Phe	Asp	Thr	Glu	Phe	Glu	Arg	Glu	Ile	Lys
							100		105			110			

Arg	Glu	Val	Phe	Lys	Ile	Val	Arg	Arg	Gly	Lys	Asp	Phe	Gly	Thr	Leu
							115		120			125			

Gly	Trp	Phe	Phe	Arg	Ala	Phe	Cys	Tyr	Ile	Ala	Ile	Phe	Phe	Tyr	Leu
							130		135			140			

Gln	Tyr	His	Trp	Val	Thr	Thr	Gly	Thr	Ser	Trp	Leu	Leu	Ala	Val	Ala
							145		150		155		160		

Tyr	Gly	Ile	Ser	Gln	Ala	Met	Ile	Gly	Met	Asn	Val	Gln	His	Asp	Ala
							165		170			175			

Asn	His	Gly	Ala	Thr	Ser	Lys	Arg	Pro	Trp	Val	Asn	Asp	Met	Leu	Gly
							180		185			190			

Leu	Gly	Ala	Asp	Phe	Ile	Gly	Gly	Ser	Lys	Trp	Leu	Trp	Gln	Glu	Gln
							195		200			205			

His	Trp	Thr	His	His	Ala	Tyr	Thr	Asn	His	Ala	Glu	Met	Asp	Pro	Asp
							210		215			220			

Ser	Phe	Gly	Ala	Glu	Pro	Met	Leu	Leu	Phe	Asn	Asp	Tyr	Pro	Leu	Asp
							225		230			235		240	

His	Pro	Ala	Arg	Thr	Trp	Leu	His	Arg	Phe	Gln	Ala	Phe	Phe	Tyr	Met
							245		250			255			

Pro	Val	Leu	Ala	Gly	Tyr	Trp	Leu	Ser	Ala	Val	Phe	Asn	Pro	Gln	Ile
							260		265			270			

Leu	Asp	Leu	Gln	Gln	Arg	Gly	Ala	Leu	Ser	Val	Gly	Ile	Arg	Leu	Asp
							275		280			285			

Asn Ala Phe Ile His Ser Arg Arg Lys Tyr Ala Val Phe Trp Arg Ala
 290 295 300
 Val Tyr Ile Ala Val Asn Val Ile Ala Pro Phe Tyr Thr Asn Ser Gly
 305 310 315 320
 Leu Glu Trp Ser Trp Arg Val Phe Gly Asn Ile Met Leu Met Gly Val
 325 330 335
 Ala Glu Ser Leu Ala Leu Ala Val Leu Phe Ser Leu Ser His Asn Phe
 340 345 350
 Glu Ser Ala Asp Arg Asp Pro Thr Ala Pro Leu Lys Lys Thr Gly Glu
 355 360 365
 Pro Val Asp Trp Phe Lys Thr Gln Val Glu Thr Ser Cys Thr Tyr Gly
 370 375 380
 Gly Phe Leu Ser Gly Cys Phe Thr Gly Gly Leu Asn Phe Gln Val Glu
 385 390 395 400
 His His Leu Phe Pro Arg Met Ser Ser Ala Trp Tyr Pro Tyr Ile Ala
 405 410 415
 Pro Lys Val Arg Glu Ile Cys Ala Lys His Gly Val His Tyr Ala Tyr
 420 425 430
 Tyr Pro Trp Ile His Gln Asn Phe Leu Ser Thr Val Arg Tyr Met His
 435 440 445
 Ala Ala Gly Thr Gly Ala Asn Trp Arg Gln Met Ala Arg Glu Asn Pro
 450 455 460
 Leu Thr Gly Arg Ala
 465

<210> 32
 <211> 26
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Polylinker

<400> 32
 gaattcggcg cgccgagctc ctgcag

26

<210> 33
 <211> 265
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Polylinker-Terminator-Polylinker

<400> 33
ccaccgcgtt gggcgccgc ctgcagtcta gaaggcctcc tgcttaatg agatatgcga 60
gacgcctatg atcgcatgat atttgcttc aattctgttgc tgcacgttgt aaaaaacctg 120
agcatgtgtt gctcagatcc ttaccgcggg ttccggttca ttctaatgaa tatatcaccc 180
gttactatcg tattttatcg aataatattc tccgttcaat ttactgatttgc tccgtcgacg 240
aattcgagct cggcgccca agctt 265

<210> 34
<211> 257
<212> DNA
<213> Artificial Sequence

<220>
<223> Polylinker-Terminator-Polylinker

<400> 34
ggatccgata tcggggccgc tagcgtaac cctgcttaaa tgagatatgc gagacgccta 60
tgatcgcatg atatttgctt tcaatttgcgtt tgcacgtt gtaaaaaaaaacc tgagcatgtg 120
tagctcagat ccttaccgccc ggtttcgggtt cattctaatg aatataatcac ccgttactat 180
cgtatttta tgaataatata tctccgttca atttactgat tgcgtcgacg cgaattcgag 240
ctcggcgccgc caagctt 257

<210> 35
<211> 257
<212> DNA
<213> Artificial Sequence

<220>
<223> Polylinker-Terminator-Polylinker

<400> 35
agatctgccc gcatcgatcc cggccatgg cctgcttaaa tgagatatgc gagacgccta 60
tgatcgcatg atatttgctt tcaatttgcgtt tgcacgtt gtaaaaaaaaacc tgagcatgtg 120
tagctcagat ccttaccgccc ggtttcgggtt cattctaatg aatataatcac ccgttactat 180
cgtatttta tgaataatata tctccgttca atttactgat tgcgtcgacg cgaattcgag 240
ctcggcgccgc caagctt 257